

Hacettepe Üniversitesi Sosyal Bilimler Enstitüsü
Felsefe Anabilim Dalı

# DANIEL DENNETT'IN ZİHİN FELSEFESİNDE GÜÇLÜ YAPAY ZEKANIN OLASILIĞI

Ayşe ÇAĞIL

Yüksek Lisans Tezi

Ankara, 2022

Daniel Dennett'ın Zihin Felsefesinde Güçlü Yapay Zekanın Olasılığı

Ayşe ÇAĞIL

Hacettepe Üniversitesi Sosyal Bilimler Enstitüsü
Felsefe Anabilim Dalı

Yüksek Lisans Tezi

Ankara, 2022

## KABUL VE ONAY

Ayşe Çağıl tarafından hazırlanan "Daniel Dennett'ın Zihin Felsefesinde Güçlü Yapay Zekanın Olasılığı" başlıklı bu çalışma, 11.4.2022 tarihinde yapılan savunma sınavı sonucunda başarılı bulunarak jürimiz tarafından Yüksek Lisans Tezi olarak kabul edilmiştir.

Doç. Dr. Çetin Türkyılmaz (Başkan)

Prof. Dr. Nazile Kalaycı (Danışman)

Doç. Dr. Aziz Fevzi Zambak (Üye)

Bu tez çalışmasında Sayın Prof. Dr. Murat Arıcı Ortak Danışman olarak görev almıştır.

Yukarıdaki imzaların adı geçen öğretim üyelerine ait olduğunu onaylarım.

Prof. Dr. Uğur Ömürgönülşen

Enstitü Müdürü

# YAYIMLAMA VE FİKRİ MÜLKİYET HAKLARI BEYANI

Enstitü tarafından onaylanan lisansüstü tezimin tamamını veya herhangi bir kısmını, basılı (kağıt) ve elektronik formatta arşivleme ve aşağıda verilen koşullarla kullanıma açma iznini Hacettepe Üniversitesine verdiğimi bildiririm. Bu izinle Üniversiteye verilen kullanım hakları dışındaki tüm fikri mülkiyet haklarım bende kalacak, tezimin tamamının ya da bir bölümünün gelecekteki çalışmalarda (makale, kitap, lisans ve patent vb.) kullanım hakları bana ait olacaktır.

Tezin kendi orijinal çalışmam olduğunu, başkalarının haklarını ihlal etmediğimi ve tezimin tek yetkili sahibi olduğumu beyan ve taahhüt ederim. Tezimde yer alan telif hakkı bulunan ve sahiplerinden yazılı izin alınarak kullanılması zorunlu metinleri yazılı izin alınarak kullandığımı ve istenildiğinde suretlerini Üniversiteye teslim etmeyi taahhüt ederim.

Yükseköğretim Kurulu tarafından yayınlanan ***"Lisansüstü Tezlerin Elektronik Ortamda Toplanması, Düzenlenmesi ve Erişime Açılmasına İlişkin Yönerge"*** kapsamında tezim aşağıda belirtilen koşullar haricince YÖK Ulusal Tez Merkezi / H.Ü. Kütüphaneleri Açık Erişim Sisteminde erişime açılır.

- o Enstitü / Fakülte yönetim kurulu kararı ile tezimin erişime açılması mezuniyet tarihimden itibaren 2 yıl ertelenmiştir. (1)
- o Enstitü / Fakülte yönetim kurulunun gerekçeli kararı ile tezimin erişime açılması mezuniyet tarihimden itibaren ….. ay ertelenmiştir. (2)
- o Tezimle ilgili gizlilik kararı verilmiştir. (3)

11/4/2022

# ETİK BEYAN

Bu çalışmadaki bütün bilgi ve belgeleri akademik kurallar çerçevesinde elde ettiğimi, görsel, işitsel ve yazılı tüm bilgi ve sonuçları bilimsel ahlak kurallarına uygun olarak sunduğumu, kullandığım verilerde herhangi bir tahrifat yapmadığımı, yararlandığım kaynaklara bilimsel normlara uygun olarak atıfta bulunduğumu, tezimin kaynak gösterilen durumlar dışında özgün olduğunu, **Prof. Dr. Nazile KALAYCI** danışmanlığında tarafımdan üretildiğini ve Hacettepe Üniversitesi Sosyal Bilimler Enstitüsü Tez Yazım Yönergesine göre yazıldığını beyan ederim.

***Dr. Ayşe ÇAĞIL***

# ÖZET

ÇAĞIL, Ayşe. Daniel Dennett'ın Zihin Felsefesinde Güçlü Yapay Zekanın Olasılığı, Yüksek Lisans Tezi, Ankara, 2022.

Bu çalışmada güçlü yapay zekanın olanaklılığı, güncel zihin modellerinin ön görüleri açısından incelenmiştir. Yapay zeka sistemleriyle ortak noktalarının olmasından dolayı, Dennett'ın zihin modeli baz alınarak, bu modelin yapısı ve bilince yaklaşımı incelenmiş, modelin güçlü yapay zekanın olanaklığını ne ölçüde desteklediği araştırılmıştır.

**Anahtar Sözcükler**

Güçlü yapay zeka, Dennett, Çoklu Taslaklar Modeli, Evrim, Bilinç, Bilincin zor problemi

# ABSTRACT

ÇAĞIL, Ayşe. Possibility of strong artificial intelligence in Daniel Dennett's philosophy of mind, Yüksek Lisans Tezi, Ankara, 2022.

In this thesis, possibility of strong artificial intelligence is investigated in the domain of recent mind models. Since Dennett's model of consciousness have common insights with artificial intelligent systems, Dennett's model of concsiousness is reviewed and possibility of strong artificial intelligence according to Dennett's model of consciousness is explored.

**Keywords**

Artificial intelligence, Possibility of strong artificial intelligence, Hard problem of consciousness, Multiple Drafts Model, Dennett, Evolution

# İÇİNDEKİLER

# GİRİŞ

Günümüzde yapay zeka sistemleri hayatın her alanına girerken yapay zekanın limitlerinin ne olduğu önemli bir soru haline gelmiştir. Yapay zeka çalışmalarının başlangıcından itibaren yapay zeka alanı öncülerinin hedefleri insan gibi zeki davranışlar sergileyebilen sistemler yapmak olmuştur. Bazı yapay zeka araştırmacıları insan gibi bilinçli makinelerin yapılmasının dahi mümkün olduğunu savunmuşturlar. Bu fikir ayrılığı, “zayıf yapay zeka” ve “güçlü yapay zeka” kavramlarını üretmiştir. İlk kez Searle tarafından öne sürülen bu kavramlardan zayıf yapay zeka, çok karmaşık problemleri çözebilen ama temelde hesap yapmaktan öteye gitmeyen sistemler için kullanılırken, güçlü yapay zeka kavramı öğrenebilen, düşünebilen, kendi farkındalığı olması hedeflenen sistemleri betimlemek için kullanılmaya başlanmıştır. Tanımı gereği güçlü bir yapay zeka sistemi bilinçli bir sistem olacaktır. Bilinç kavramı ise henüz anlaşılabilmiş bir kavram değildir. Bilincin nörobilim araştırmalarına girmesi oldukça yeni gelişmedir. Chalmers nörobilim araştırmalarının çözmeye çalıştığı sorunları bilincin kolay problemleri olarak tanımlamıştır. Duyusal uyaranların ayırt edilmesi, beynin bilgiyi nasıl entegre ettiği, uzun süreli bellek oluşumu, düşünme ve karar verme mekanizmaları, zihinsel durumları bildirme, dikkat mekanizmaları işlevsel bir bakış açısıyla belli fiziksel bağlantılar ve mekanizmalarla açıklanabilecek problemlerdir. Chalmers bilincin zor problemini ise neden öznel deneyimlerimiz olduğu, öznel bilinç içeriklerimizin fiziksel bir sistem tarafından nasıl yaratıldığı sorularıyla tanımlamıştır. (Chalmers, 1995, s: 202) Chalmers bilincin fiziksel sistemlerle açıklanamayacağı görüşünü savunurken Dennett insan bilincinin fiziksel sistemler tarafından yaratıldığını, aynı şekilde belli koşullara uyan sistemlerin bilinçli olmasının mümkün olduğunu savunmuştur. Dennett’ın savunmasının temeli bilincin bir işlevi yerine getirmek üzere evrimin bir aşamasında çıkmış olması ve bilincin beynin karmaşıklığı sonucu zayıf beliren bir özellik oluşudur.

Bilgisayar bilimlerinin öncüsü Alan Turing’in bir matematik problemi çözümü için tasarladığı hipotetik makine günümüz bilgisayarlarının temelini atmış olan otomatik bir makinedir. Turing makinesi giriş sembollerini belirlenmiş sonlu kural dizileriyle manipüle ediyor ve giriş sembollerine göre çalışan sonlu kural dizisi her sembol için sonsuz kez kullanılabiliyordu. (Turing, 1936, s: 233-235) Bu kurallar dizisi günümüzde

algoritma olarak bilinir. Von Neuman, Turing'in hipotetik makinesini gerçek bir makineye dönüştürerek ilk bilgisayarları inşa etmiştir. Günümüzde kullanılan bilgisayarlar hala Von Neuman mimarisini kullanarak çalışmaktadır. 1950 yılında Turing, bilgisayarların her türlü hesabı yapabileceğini, hatta bilgisayarların insan zekasını taklit edebileceğini öne sürmüş ve bir bilgisayarın insanın zihinsel yetilerini taklit etme kabiliyetini sınamak için bir taklit oyunu tasarlamıştır. (Turing, 1950, s: 433) Günümüzde Turing testi olarak bilinen test, bir bilgisayarın bir insanın düşünme sürecini taklit edip edemeyeceğini sınayan bir testtir.

Turing makinesinin önemi insanların yaptığı hesapların bir makine tarafından da yapılabilir olduğunun anlaşılmasıdır. 1940'lı yıllara dek bilişsel bilimler Descartes'ın töz düalizm etkisinde zihinsel özelliklerin fiziksel olmayan özellikler olduğu düşüme eğiliminde araştırmalarını sürdürmüştür. Bilişsel bilimlerde hakim görüş olan davranışçılık zihni bir kara kutu olarak kabul ederek bilişsel bilimleri davranış girdi ve çıktılarına indirgemeye çalışmıştır. Ancak 1950'li yıllara gelindiğinde bazı felsefeciler zihinsel özelliklerin beyin süreçleri olduğunu ve söz konusu zihinsel süreçlerin davranışların nedeni olduğunu savunarak özdeşlik kuramını ortaya atmıştır. Özdeşlik kuramı zihinsel süreçleri beynin fiziksel süreçleriyle özdeşleştirir. Putnam ise özdeşlik kuramına karşı çıkarak sinir sistemi insanınkinden çok farklı olan ahtapotların acı çekmesinin düşünülebilir olduğunu, bu taktirde zihinsel durumların beynin fiziksel durumlarıyla özdeş olamayacağını, zihinsel durumların çoklu gerçekleştirilebilir olduğunu savunmuştur. (Putnam, 1967) Zihinsel durumların çoklu gerçekleştirilebilir olması ise işlevselcilik görüşünün temel düşüncesidir. İşlevselcilik beynin zihinsel durumları gerçekleştiren bir organ olduğu, zihinsel durumların ise işlevsel özellikler olup, zihinsel durumların diğer zihinsel durumların, algıların ve davranışların nedensel ilişkileriyle belirlendiğini savunur. Putnam tarafından öne sürülen makine işlevselliği insan zihninin Turing makinesi gibi sonlu durumlardan ve bu durumlar üstüne tanımlanan hesaplamalardan meydana geldiğini iddia etmiş, girdi ve çıktılar arasındaki gerçekleştiricilerin ne olduğunu bilmediğimizi varsayarmıştır. (Putnam, 1967) Fodor ise zihinsel kavramları oynadıkları roller üstünden tanımlar. Fodor halk psikolojisinin üst düzey bir bilim olduğunu zihin modelinin üst seviyedeki özellikler üstünden tasarlanması gerektiğini savunmuştur. Halk psikolojisinin kavramları olan inanç, istek ve arzular gibi zihinsel olaylar önermesel tutumlardır. Fodor, zihnin semantik anlamları olan sembolleri manipüle eden bir Turing makinesine eşdeğer olduğunu ve kurallar listesinin önermesel tutumlarla ifade edilebilir olduğunu iddia etmiştir. (Fodor, 1994, s: 7) Düşüncenin dili teorisi tüm insanların zihninde ortak ve

kurallı bir yapı olduğunu ve söz konusu yapının konuşulan dilin yapısıyla aynı yapıya sahip olması gerektiğini belirtir. Fodor önermesel tutumların gerçekliğini ve zihnin üst seviye bir yapı olduğunu kabul ederek zihinsel temsillerimizin güçlü belirmiş özellikler olduğunu, zihinsel temsilerin özneye sunuluşu olarak tanımladığımız *qualia*'nın fiziksel olmayan bir var olanı tanımladığını kabul eder.

Dennett, Fodor'un zihin kuramının merkezindeki inanç, arzu, istek gibi kavramların zihni anlamamıza yardımcı olan araçsal kavramlar olduğunu ve üst düzey bilimlerle alt düzey bilimlerin arasında bir kopukluk olmadığını düşünür. Dennett'a göre yönelimsel bir duruşu benimseyerek inanç, arzu, istek gibi kavramları organizmalara atfeden bizizdir. Halk psikolojisi bu kavramları organizmalara söz konusu kavramların belli davranışları öngörme gücü olduğu için atfeder ancak bu kavramların gerçek kavramlar olduğunu savunmamız için elimizde bir neden yoktur. Dennett yönelimsel duruşun yanı sıra fiziksel duruş ve tasarım duruşu olarak iki duruş daha tanımlar. Fiziksel duruş bir sistemin hareketini fizik yasaları üstünden anlamaya ve öngörmeye çalışır. Tasarımsal duruş ise bir sistemin tasarımındaki belli bir amacı yerine getirmek üzere tasarlanmış olduğunu düşünerek sistemin hareket veya davranışları anlayıp, öngörür. (Dennett, 1987, s: 2) Bir sisteme hangi durumu atfedeceğimizi hedeflerimiz doğrultusunda seçen her zaman bizizdir. Bilim sistemlerin davranışlarını ön görmeye çalışır, temel bilimler fiziksel duruşla sistemlerin davranışlarını matematiksel bir kesinlikle öngörebilmektedirler. Ancak sistemler karmaşıklaştığında hesaplama zorluğundan veya imkansızlığından dolayı fiziksel duruş yerine tasarımsal ve yönelimsel duruşu seçebiliriz. Yönelimsel duruşla açıkladığımız sistemler her zaman için yönelimsel duruşa sahip olan sistemler olmayabilir, hatta bir sistemin gerçekten inanç ve arzulara sahip olup olmadığını anlamamızı sağlayacak bilimsel bir yöntem yoktur. (Dennett, 1987, s: 8) Dennett gerçek anlamda yönelimsel duruşa sahip olan sistemlerin ikinci dereceden yönelimsel duruşa sahip olan sistemler olduğunu iddia eder. Dolayısıyla Dennett inanç, arzu gibi zihinsel kavramların epistemolojik olarak kullanışlı araçlar olduğunu belirterek bu kavramların ontolojik temellerini dışlamış, zihin ve yapay zeka araştırmalarını tamamen epistemolojik bir zemine indirgemiştir.

Dennett indirgemeci bir filozof olarak zihinsel özelliklerin beyinin özelliklerinden belirdiğini ve fiziksel bir duruşla en nihayetinde zihinsel özelliklerin nöronlar ile açıklanabileceğini savunur. Beyin karmaşıklığından dolayı çok düzeyli bir yapıdır ve her düzeyde değişik düzenlilikler göstermektedir. Dolayısıyla Dennett homunkular işlevselliğin zihni incelemek için en elverişli işlevsel bakış açısı olduğunu düşünür.

Dennett zihne ait kavramları araçsal kavramlar olarak tanımlayarak bu kavramların ontolojisinin incelenmesinin mevzu bahis olmadığını gösterip güçlü yapay zekanın olanaklılığının zemini hazırlamış ve güçlü bir yapay zeka yaratmanın olanaklılığının epistemolojik bir problemin çözülebilme olanaklılığına indirgemiştir. Yapay zeka çalışmaları en temelde zihnin çoklu gerçekleştirilebilir olduğu kabul ederek tartışmalarını yapay sistemlerin hangi tip işlevsel model üstüne kurulabileceği çerçevesine tartışırlar. Yapay zeka araştırmacıları da güçlü yapay zekanın olanaklılığını epistemolojik bir sorun olarak görme eğilimindedir.

Yapay zeka araştırmalarının tam teşekkülü bir araştırma alanına dönüşmesi ise 1950'li yıllara tekabül eder. 1956 yılında Minsky yapay zekanın sahip olması gereken becerileri görüntü tanıma, öğrenme, problem çözme, karar verme, plan yapabilme olarak tayin ederek, bu becerilerin matematiksel olarak nasıl ifade edilebileceğini gösteren taslaklar oluşturmuştur. Minsky "Yapay Zekaya Doğru" makalesinde, eğer bir varlık, hipotetik bir soruyu, deney yapmadan cevaplama yetisine sahipse cevabın ancak varlığın içindeki bir modelden gelebileceğini, bu nedenle konulması gereken hedefin içsel bir modele göre davranan bir makine inşa etmek olduğunu belirtmiştir. (Minsky, 1961, s: 28)

İlk başlangıç tarihinden günümüze yapay zeka çalışmaları iki odağa ayrılmıştır. Yapay sinir ağları araştırmaları doğrudan beynin biyolojik yapısını temel alarak, nöronların ve nöron etkileşimlerinin çalışma prensiplerini taklit eden makineler tasarlamaya yönelirken, günümüzde eski iyi yapay zeka olarak bilinen ve sembolik mantığa dayalı yapay zeka araştırmalarıysa insan zihninin çalışma kurallarını taklit eden makineler tasarlamaya yönelmiştir. Beynin biyolojik yapısından esinlenen yapay zeka çalışmalarının temelinde nöron hücrelerinin etkileşimleriyle insan zihninin tüm yetilerinin açıklanabileceği görüşü yatar ki bu görüş günümüzde "bağlantıcılık" olarak adlandırılır. Sembolik mantığa dayalı yapay zeka çalışmaları Fodor'un temsili zihin kuramındaki gibi zihnin üst düzey özellikleriyle ilgilenmiş ve bilinçli düşünceyi taklit eden mimariler üstünden zeki davranışlar sergileyen sistemler tasarlamıştırlar. Bu çalışmalar fiziksel simge hipoteziyle temellenmiştirler, fiziksel simge hipotezi sembol yapılarının dönüştürülmesiyle çalışan sistemlerin zeki eylemler için gerekli ve yeterli imkanlara sahip olduğunu öne sürer.

Bağlantıcılık temelli yapay zeka araştırmalarıysa insan beyninin yapısını ve nöronların birbirileriyle etkileşimlerini taklit ederek insan gibi düşünen makineler tasarlamaya yönelmiştir. İlk yapay sinir ağı olan perceptronun temelinde nöron kâküllüsü ve

biyolojik beyinlerde öğrenmenin nöron bağlantıları arasındaki fiziksel ve kimyasal değişikliklerle gerçekleştiğini öne süren Hebb öğrenme kanunu bulunmaktadır. Beynin yapısı ve işlevleri anlaşıldıkça yapay sinir ağları araştırmacıları nörobilim araştırmalarının bulgularını kullanarak yapay sinir ağı mimarileri oluşturmuştur. Örneğin örüntü tanıyan evrişimli sinir ağları Hubel ve Wiesel'in tarafından birincil görme sisteminde tespit edilen özellik seçici nöronlardan esinlenerek tasarlanmış bir mimaridir. (Fukushima, 1980, s: 193-202) İnsan beyni evrende bilinen en karmaşık sistemdir; insan beyninin yaklaşık yüz milyar nörondan oluştuğu ve her nöronun diğer nöronlarla yaklaşık bin ila on bin bağlantı yaptığı bilinmektedir. Modüler bir yapıda olan insan beyni, aynı zamanda bütün bir sistem olarak çalışır. Değişik modüller, değişik işlevleri hesaplarken tüm diğer modüller ile iletişim halindedir. İnsan beyni paralel dağıtılmış, her biri eşzamanlı olarak birçok işlevi yerine getiren nöronal devrelerden oluşur. İnsan beyninin ve genel olarak omurgalı beyinlerinin en önemli özelliği ise bir geleceği tahmin etme makinesi gibi işlev görmeleridir. Beynimiz her an bir sonraki anda ne olacağını hesaplamaktadır. Beyin dışsal duyuları değişik modüllerinde hiyerarşik bir yapıda adım adım işlerken her adımda üst seviye modüllerle döngüsel bir iletişim kurmaktadır. Beyin duyusal bir algıyı değerlendirirken, duyusal algıdaki ipuçlarını önceden oluşturmuş olduğu temsillerle kıyaslayarak en olası ve doğru çıktıyı tahmin etmeye çalışır. Böylece hesaplama süresi minimize edildiği gibi duyusal verideki eksikler tamamlanır. Görüntüleme teknikleri beyinde alt modüllerden üst modüllere olan bilgi akışından çok üst modüllerden alt modüllere bilgi akışı olduğunu tespit etmiştir. Beynin üst düzey modüllerinden alt düzey modüllerine bilgi akışı “yukarıdan aşağıya işleme” olarak adlandırılırlar. Tekrarlayan sinir ağları da iç temsillerini dinamik olarak oluşturur ve yeni duyusal bilgiyi iç temsilleriyle yorumlar. Bu ağlar aynen beynin duyusal algıyı işleyen kısımları gibi her yeni bilgide içsel temsillerini günceller. Beynin dış duyusal bilgiyi işleyerek duyusal bilgideki ipuçlarından olası en doğru çıkarımı yapmaya çalıştığını öne süren modellere genel olarak Bayesci modeller denmektedir. Friston, Bayesci modellerin nasıl iş gördüğünü “serbest enerji prensibi” ile fiziksel bir temele oturtmuştur. Serbest enerji prensibi tek bir hücreden insan beyni gibi en karışık biyolojik sistemlere uygulanabilen bir prensiptir ve kendi kendini organize eden tüm biyolojik yapıların kendi bütünlüklerini korumak için dış dünya ile kendi dış dünya dair iç temsilleri arasındaki ön görü farkını en aza indirmeye çalışır gibi bir davranış sergilediklerini iddia eder. Serbest enerji prensibine göre beyin dış duyusal bilgi ile iç temsilleri arasındaki farkı, yani serbest enerjisini minimize etmeye çalışmaktadır. (Friston, 2012, s: 1232-1233) Dolayısıyla

Friston'a göre tekrarlayan sinir ağları, beynin bir modelidir. Tekrarlayan sinir ağlarının dış dünyanın temsillerini içsel modellerinde yaratmalarının en çarpıcı özelliği beynin bilinçli içeriğinin yukarıdan aşağı işlemeyle alakalı bir kavram olduğu düşüncesinden kaynaklanmaktadır. Beynimiz dış dünyayı içsel temsilleri aracılığıyla kendi yapısında oluşturur, beynimizdeki temsiller dış dünyanın özelliklerini içerirler. Temsilciliğe göre bir şeyi bilinçli olarak deneyimlememizin, o şeyin zihinsel temsili ile ilgilidir. Tüm zihinsel temsiller bilinçli değildir ancak bilinçli her deneyim beynin içsel temsillerini içerir. Eğer ki beynimizde yukarıdan aşağı bir işleme söz konusu ise, yukarıdan aşağı işlemeyi meydana getiren zihinsel temsilin bilinçli bir içerik oluşturması muhtemeldir. En nihayetinde zihnin bilinçli içerikleri her zaman için ikinci dereceden bir temsildir; dış dünyanın yaratılmış içsel temsilinin tekrardan temsili.

Bilinçli deneyim ahlak, özgür irade, aşk, acı gibi insanlığın değer verdiği ve değerlerini yarattığı kavramlarla yakından alakalı bir kavram olduğu halde bilim bilinci incelemekten yakın zamana dek kaçınmıştır. Francis Crick, 1990 yılında kaleme almış olduğu *Şaşırtan Varsayım* kitabında nörobilim dalında çalışan bilim insanlarının bilinci merak etmekte son derece isteksiz olduklarını vurgular. Crick, bilimin bilinci araştırmaktan kaçınmasının nedenleri arasında davranışçı ekolün bilinci içine girilmez bir kara kutu olarak göstermiş olmasının yanı sıra bilinci incelemek adına doğru bilimsel yöntemlerin belirlenmemiş olması olduğunu belirtmiştir. Crick, bilincin tamamen beynin çalışma prensiplerinin bir ürünü olduğunu belirtir. Crick bu varsayımı "şaşırtan varsayım" olarak adlandırır. "Şaşırtan varsayım: siz, neşeleriniz, üzüntüleriniz, anılarınız, ihtiraslarınız, benlik ve özgür irade duygularınızla, aslında çok sayıda nöron ve bunlarla ilişkili moleküllerin bir arada davranışından ibaretsiniz." (Crick, 1997, s: 1) Crick'in ikinci varsayımı ise acı veya mutluluk hissetmek, görsel veya duyumsal farkındalık gibi bilincin çeşitli yönlerinin ortak bir işleyişin ya da düzenin sonucu olduğudur. Dolayısıyla Crick, bilinçli bir deneyimin, örneğin bir görselin bilinçli biçimde idrak edilmesinin nasıl gerçekleştiğini ya da beyinde görsel bilinçli deneyimin hangi nöronal sistemlerin hangi dinamikleriyle ortaya çıktığını bulduğu taktirde genel olarak tüm bilinçli deneylerimin açıklanabileceğini savunmuştur. (Crick, 1997, s: 23) İki temel varsayıma ilave olarak Crick, bilinç için bilimsel bir araştırma zemini belirlerken bilincin kesin bir tanımını yapmaktan kaçınmıştır. Crick yapılacak bir tanımın bilinç araştırmalarını kısıtlama ve yanlış bir yola yönlendirme olasılığı olacağını belirtmiştir. Crick ve Koch, bilinci bilimsel bir metotla çalışmanın en basit yolunun beyinde bilinçli deneyime neden olan minimal nöron bağlantılarının bulunması olduğunu ve bilinçli deneyimin minimal nöronal

bağlantılarının görsel sistemdeki bilinçli deneyimi inceleyerek bulunabileceğini öne sürmüşlerdir. (Crick, Koch, 1990, s: 266-267)

Nörobilim araştırmacıları bilişsel bilimlerin birikimiyle bilincin nöronal korelasyonları araştırmalarından elde ettikleri verileri bütünleştirerek bilince dair modeller oluşturmuştur. Bu modellerden Baars tarafından oluşturulan kapsamlı çalışma alanı modeli günümüzde en geçerli modellerden biridir. Baars, belli bir bilgi parçasının farkında oluşumuzun, dolayısıyla belli bir bilgi parçasının bilinçli olarak algılanmasının, söz konusu olan bilgi parçasının diğer bilişsel mekanizmaları tarafından kullanılabilir hale getirdiğini savunarak bilince işlevsel bir rol affetmiştir. Bilinçli zihnin sınırlı kapasitesine dikkat çeken Baars, beynin bilinçsiz işlem kapasitesinin genişliğiyle kıyaslandığında bilinçli kapasitenin ilk bakışta organizma için bir dezavantaj gibi gözüktüğünü ancak bilincin evrimde bir işleve hizmet etmek amacıyla ortaya çıkmış olması gerektiğini belirtir. Bilincin sağladığı avantaj beynin birçok modülünün yaptığı hesaplamaları birleştirmek ve üst düzey hesaplamalar için söz konusu bilgiyi kullanılır şekilde temsil etmektir. Bilginin bütünsel olarak kullanılabilirliği, tam da öznel olarak yaşadığımız bilinç halidir. (Deheane, 2018, s: 218)

Kapsamlı çalışma alanı modeli ve nörobilim alanındaki diğer bilinç modelleri nöronal mekanizmalardan ya da ampirik gözlemlerden bilinçli deneyimin yapısına ulaşmaya çalışmaktadır. 2004 yılında Guiliano Tonini farklı bir yöntemle bilinçli deneyimlerin özellikleriyle başlayarak bilincin nöronal mekanizmalarına ulaşmayı hedefleyen entegre bilgi teorisini yayınlamıştır. Entegre bilgi teorisi bilincin entegre edilmiş, olduğundan daha küçük parçalara indirgenemez bir bilgi olduğu sonucuna varmıştır. Entegre bilgi teorisine göre tekrarlayan sinir ağlarında bilincin varlığından söz etmek mümkündür.

Tekrarlayan yapay ağların dış dünyanın temsillerinden kendi iç temsillerini oluşturmaları ve iç temsillerini dinamik olarak güncellemeleri yapay sinir ağlarının en önemli özelliği olmuştur. Ancak biyolojik beyinleri yapay sinir ağlarından ayıran en önemli özelliklerinden biri dış dünyadaki bilgilerin hangilerinin organizma için önemli, hangilerinin organizma için önemsiz olduğunu ayırt edebilmeleridir. Biyolojik beyinler dış dünyanın temsillerini oluştururken sadece dış dünyanın organizmayı ilgilendiren durumları ile uğraşmaktadırlar. Bilinç problemine evrimsel bir açıdan yaklaşan Damasio, beyin korteksinin hatta sinir sisteminin ortaya çıkmasından önce basit organizmaların hayatta kalmasını sağlayan mekanizmalar geliştirmiş olduğunu

belirmiştir. (Damasio, 2020, s: 27) Bu mekanizmalar organizmalar karmaşıklaştıkça duygu mekanizmalarına ve belli bir düzeyde gelişmiş sinir sistemi olan canlılarda hislere evirilmişlerdir. Damasio, bilincin duygulardan bağımsız araştırılmaması gerektiğini savunur. Beyinler bilinçli zihinleri korteks düzeyindeki değil, beyin sapı düzeyindeki değer sistemleri aracılığıyla oluşturmaya başlar. Çoğu duygusal tepki beyin sapında şekillenir. (Damasio, 2020, s: 23) Beyin sapının üst kısımlarında, aralarında hipotalamusun da bulunduğu birçok çekirdek yapı bulunmaktadır. Beyin sapı çevresindeki tüm çekirdeklerin nöronları dağınık bir şekilde beynin tamamına dağılmışlardır. Bu çekirdeklerdeki nöronlar, ağrı, ani ses, ani ışık gibi beklenmedik durumlarda ve vücudundaki beklenmedik değişikliklerde ateşlenerek beyinde nöro-modülatörlerin salgılanmasına neden olurlar. Nöro-modülatörler sinapsların hem gücünü hem de bağlantılarını değiştirme gücüne sahip kimyasallardır. Bu nedenle beyin sapı çevresindeki çekirdek sistemlerine değer sistemleri denilmektedir. (Edelman, Tononi, 2019, s: 71) İnsanlarda, hislerin tamamen rasyonel bir mekanizma olduğu düşünülen karar verme mekanizmasıyla ilişkisini inceleyen Damasio, önceden normal olan bireylerin belli beyin bölgeleri hasar gördüğünde bu hastaların toplumsal ve finansal hayatlarını yönlendirme becerilerini kaybettiğini gözlemlemiştir. (Damasio, 2018, s: 144) Eylemi ödül veya ceza duygularıyla bağdaştırmaktan sorumlu bölgelerde meydana gelen hasarlar rasyonel bir mekanizma olan karar verme mekanizmasından yoksun kalmaktadır. Damasio duygulardan evrimleşmiş hislerin bilinçli deneyimle sıkı bir etkileşim içinde olduğunu düşünür. Hisleri bilinç olmaksızın, bilinci ise hisler olmaksızın anlamak mümkün değildir. Damasio bilinçli bir zihnin temelinde bedenin olduğunu vurgular. Beyin içinde bulunduğu vücuda hizmet eder ve beynin esas fonksiyonu karmaşık organizmaların dengeleşimini sağlamaktır. Beyne devamlı olarak vücudun içsel durumları ile ilgili bilgi iletilmektedir. Beyinde aynen duyusal girdiler ile şekillenmiş temsiller olduğu gibi vücudun içsel durumlarının nöronal temsilleri de bulunmaktadır. Vücudun içsel mekanizmalarında oluşan durumlar, acıkmak, susamak, halsiz hissetmek gibi *qualia* boyutunda hislere denk gelir. Bu mekanizmalarını yapay sinir ağlarına uygulanması ise pekiştirmeli öğrenme alanını doğurmuştur. Pekiştirmeli öğrenme en basit tanımıyla deneme yanılma yoluyla öğrenmedir. Pekiştirmeli öğrenme bir amaç doğrultusunda çevre ile etkileşime girerek zaman içinde öğrenen ve öğrendikleri ile bir probleme yaklaşımlar geliştirebilen özerk ajanlar tasarlamayı hedeflemiştir. Günümüzde kapalı çevrelerde, örneğin atari oyunlarında pekiştirmeli öğrenme algoritmaları insan üstü yetenekler sergilemeyi başarmıştır.

Günümüzde güçlü yapay zekanın bağlantıcılık temelli modeller ile daha mümkün olduğu görüşü ağırlık kazanmış olsa da bağlantıcılık temelli modellerin insan zihninin birçok yetisini taklit edemeyeceğini savunan görüşler bulunmaktadır. İnsan zihninin yetilerinin yapay sinir ağlarıyla taklit edilemeyeceğini belirterek bağlantıcılığı eleştiren görüşlerin temelinde zihinsel hesaplamaların ancak sembolik bir yapıyla açıklanabileceği düşüncesi yatmaktadır. Chomsky insan dilinin kurallı bir yapısı olduğunu göstermiş, Fodor ise dilin yapısının tüm düşünce süreçlerindeki yapı ile aynı olduğunu ve tüm insanların evrensel bir zihin yapısına sahip olduğunu iddia etmiştir. Bağlantıcılığı savunanlar ise yapay ağların alt-sembollerden sembolik ifadeleri oluşturabildiğini savunmuştur. Smolensky zihnin işleyişinin üç temel seviyede, nöronal seviye, alt-sembolik seviye ve sembolik seviye anlaşılabileceğini öne sürmüştür. Alt sembolik seviye nöronlar ve nöronların aralarında kurduğu dinamik ilişkileriyle ifade edilebilir. Alt-sembolik seviyedeki ince yapılı dağınık dağıtılmış nöronal yapılar ise kavramların, sembollerin ve sembol dönüşümlerinin yapıtaşlarıdır. (Smolensky, 1988, s: 3) Çok boyutlu bağlantı uzayında bir grup nöronun ifade ettiği temsil, tüm nöronlar arasındaki bağlantı ağırlıklarından oluşan çok boyutlu bir vektör ile ifade edilebilir. Söz konusu vektör ise, örneğin söz konusu nöron kümesi bir nesneyi temsil ediyorsa, nesneyi tanımlayan sembole denk gelir. (Smolensky, 1988, s: 17) Smolensky alt-sembolik seviyede bilginin, sembolik seviyedeki gibi açık bir temsili olmadığını ancak bu bilginin belli bir şekilde ifade edildiğinde sembolleri ve semboller arasındaki kuralları oluşturabileceğini savunmuştur. Zaman içindeki verilerden çalışan bir bellek oluşturacak şekilde tasarlanan tekrarlayan sinir ağlarının sadece girdi verilerinden yola çıkarak dilbilgisi kurallarını hatasız olarak öğrenebildiği ortaya çıkması bu görüşü desteklemiştir. Ancak Fodor ve Pylyshyn bağlantıcılık temelli modellerin sadece verinin istatistiki derlemesi sayesinde dağıtılmış temsiller ve bu temsillere dair bir hafıza oluşturabildiğini ancak sembol sistemleri tarafından oluşturulabilen ve düşüncenin sahip olduğu üretkenlik, sistematiklik, birleşimsellik ve çıkarımsal tutarlılık özelliklerini yapılarında bulundurmalarının mümkün olmadığını savunmuştur. (Fodor, Pylyshyn 1988, s: 14) İnsan düşüncesinin sistematikliği insan zihninin belirli düşüncelerden yola çıkarak kurduğu bağlantılarla verili olmayan bir düşünceyi yaratma yetisidir, örneğin John Mary'i seviyor düşüncesinden yola çıkarak Mary'nin de John'u sevdiği düşüncesinin düşünebilir olması. (Fodor, Pylyshyn, 1988, s: 14) Yapay sinir ağlarının iki farklı olay arasında sistematik bir bağlantı kurma yetisine, bir başka ifadeyle belirli bir dağılımın ötesinde genelleme yetisine sahip değildirler. Genelleme yetisi iki farklı nöronal temsil arasında hiç kurulmamış bir ilişki

kurarak bir nöronal temsilin nöron bağlantı ağırlıklarına içkin kuralları tamamen farklı bir nöron örüntüsüne uygulanmasını içermek zorundadır.

Yapay sistemlerin insan zihninin yetilerini ne şekilde temsil edebileceği tartışmasının ötesinde bazı filozoflar yapay sistemlerin insan zihninin yetilerini taklit edebilmesinin mümkün olmadığını, bu sistemlerin karmaşık hesaplar yapmaktan öteye geçemeyeceklerini savunmuşturlar. Örneğin Searle Çince odası düşünce deneyinde Turing testini geçebilecek şekilde tasarlanmış doğru girdilere doğru çıktılar üretme kapasitesi olan bir sistemin semantikten yoksun olduğunu göstermiştir. Searle'ün Çince odası düşünce deneyi sembolik sistemler göz önünde bulundurularak tasarlanmış bir düşünce deneyidir ve Searle düşünce deneyinde esasında bilgisayar bilimlerindeki fiziksel simgenin temellenmesi problemini vurgulamıştır. Sembolik sistemler semboller üstünden hesaplamalar yaparken yapılan hesaplar sembollerin içeriklerinden bağımsızdır. Sembolik sistemlerin temelindeki fiziksel simge hipotezi ise kullanılan sembollerin semantik anlamları içerdiğini kabul eder ancak sembolik sistemlerde sembollerin semantik anlamlarla ilişkisini kurmak için başarılı bir algoritma geliştirilememiştir. Yapay sistemlerin hesap yapmanın ötesine geçemeyeceğini savunan bir başka filozof ise Dreyfus'tur. Dreyfus Heidegger felsefesinden yola çıkarak, dünyada tek başına anlam taşıyan varlıklar olmadığını, varlıkların anlamlarını diğer varlıklarla girdikleri ilişkilerin belirlediğini ve anlamın her zaman teknik bilgiden fazlası olduğunu vurgulamıştır. Dreyfus, insan davranışlarının türlü yaşantı ve deneyimlerle edinilmiş türlü bilgilerin sağduyuya dayalı bir birleşimi olduğu, insanların ancak dünyanın içinde var olarak dünyayı anlamlandırdığını, dolayısıyla belli kurallarla bir zihin yaratılamayacağını düşünerek fenomenolojik öznenin zihinsel temsillerinin çeşitliliğine dikkat çekmiştir. İnsan zihninin yapısı bir şekilde biçimsel formüllere indirgense dahi insanın yaşam boyu kazandığı becerilerin sayısal verilere indirgenmesi imkansız olacaktır. (Dreyfus, 1972) Bilinci oluşturan en önemli özellik fenomenal deneyimlerdir. Vücudumuza ve dış dünyaya dair duyumlarımız veya duygularımız öznel deneyimlerdir. Öznel deneyimler, öznel deneyimi deneyimleyen özne için o özne gibi olmaktır. (Nagel, 1974, s: 147) Nagel, "Yarasa Olmak Nasıl Bir Şeydir" başlıklı makalesinde birinci tekil şahıs deneyimlerinin üçüncü şahıslar tarafından erişilmez olduğunu savunmuştur. Bir yarasa olmanın, sonar sistemiyle çevreyi algılamanın nasıl bir deneyim olduğunu asla anlayamayız. Nagel, bilimin de birinci şahsın deneyimlerine ulaşamayacağını, bilimin tarafsız, üçüncü şahsın bakış açısına sahip olduğunu belirtmiştir. Nagel'e göre bir yarasanın

beyninin tüm işlevlerini çözebilsek bile yarasanın öznel deneyimlerini öğrenme şansımız olmayacaktır.

Dennett, güçlü yapay zekanın olanaklı olduğunu savunan bir filozof olarak öncelikle fenomenolojinin bilinci incelemek için doğru bir yöntem olmadığını savunmuştur. Dennett, Dreyfus'un belirttiği problemin, açıkça tanımlanmamış olsa da "çerçeve problemi" olduğunu belirtir. (Dennett, 1984, s: 184) Bir sistem çevre ile etkileşime girdiğinde sistemin içsel aksiyomlarından hangilerinin değişip hangilerinin sabit tutulması gerektiğinin belirlenmesini çerçeve problemi olarak tanımlanmıştır. Dennett çerçeve probleminin fenomenolojik bir problemden daha çok bir ajanın çevresiyle ilişkisinin belirlendiği derin epistemolojik bir problem olduğunu belirtmiştir. Güçlü yapay zekanın olanaklılığı epistemolojik bir problem olan çerçeve probleminin çözülmesiyle alakalıdır. Çerçeve problemi ise oldukça geniş bir problemdir. Çerçeve problemi ilk olarak belirli bir durumda hangi bilgilerin o durumla alakalı, hangi bilgilerin durumla alakasız olduğunun tayin edilmesidir. Zeki bir ajanın bir durumda ancak o durumla ilgili bilgileri işleme koyması beklenir. Durumla ilgili bilgileri belirleyen bir ajanın hareketlerinin sonuçlarını tahmin etme yeteneği olması, hareketin sonucunda dış dünyanın temsillerinin iç temsillerini nasıl değiştireceği tahmin edebilmesi gerekir. Bir başka sorunsa dünyanın robotun hareketlerinden de bağımsız olarak devamlı değişmesidir. Dolayısıyla robotun dış dünyaya ait içsel temsilleri dinamik olarak dünya ile birlikte değişebilmelidir. Aynı zamanda çerçeve problemi bir zeki ajanın neleri bilmesi gerektiğini ve bu bilgileri nasıl elde edeceğiyle de ilgilidir. (Dennett, 1984)

Çerçeve problemi sadece yapay sistemlerin karşılaştığı bir sorun değildir. Herhangi bir hesaplamaya dayanan bilişsel modelde çerçeve problemine nasıl bir çözüm önerdiğini açıklamalıdır. Kapsamlı çalışma alanı modeli çerçeve problemini duygularla birleştirerek çözmektedir. Ancak biyolojik organizmalar çerçeve problemi çevreyle etkileşim içinde evrimleşerek çözmüştürler, yapay sistemler ise çerçeve problemini tersine mühendislik yaparak çözmek durumundadır. Dennett, yapay bir sistemin bilinçli olmasından veya bilinçle yakından alakalı yetileri olmasından daha çok sistemin çerçeve problemini çözecek yetileri olup, olmadığıyla ilgilenmekte ve ontolojik tartışmaları çerçeve probleminin dışında tutmaktadır. Ancak nörobilim ve yapay zeka araştırmaları büyük bir kısmı bilincin anlaşılmasının çerçeve probleminin çözümüne ulaşmak için gerekli olduğunu düşünmektedir. Nörobilim araştırmaları bilincin ne olduğunu ve bilinçli mekanizmaların tam olarak nasıl işlediğini çözmemiş olsa da nörobilim alanındaki araştırmacılar, bilinç probleminin beynin anatomik ve

fonksiyonel yapıları anlaşıldıkça çözüme kavuşacağına inanmaktadır. En nihayetinde nörobilim bilincin temelinde beynin belli özellikleri olduğunu veya zihnin beynin karmaşıklığı sonucu beliren bir özellik olduğunu varsaymaktadır.

David Chalmers bilimin duyusal uyaranların ayırt edilmesi, beynin bilgiyi nasıl entegre ettiği, uzun süreli bellek oluşumu, düşünme ve karar verme mekanizmaları, zihinsel durumları bildirme, dikkat mekanizmaları gibi ancak bilincin belli özelliklerini açıklayabileceğini ancak bilincin tam bir açıklamasının bilincin zor problemi olarak tanımlanan neden öznel deneyimlerimiz olduğu, öznel bilinç içeriklerimizin fiziksel bir sistem tarafından nasıl yaratıldığı sorunlarının yanıtlarını da içermesi gerektiğini savunmuştur. (Chalmers, 1995, s: 202) Chalmers, fenomenal deneyimlerimizin, sahip olduğumuz *qualia* uzayının hayatımıza anlam katan bir özellik olduğunu belirtir. Ancak fenomenal deneyimlerimizin beynimizin fiziksel özellikleriyle bağlantılı olsa bile beynin fiziksel özelliklerinden yola çıkarak anlaşılamayacağını savunur. Dennett ise zor problemin gerçek olmayan bir problem olduğunda ısrarcıdır. Dennett'a göre zor problemin temelinde yönelimsel duruşun seçilmesi vardır. Bilince fiziksel bir duruş edinerek yaklaştığımızda bilincin fiziksel bir sistemin karmaşıklığından dolayı beliren, işlevsel bir özellik olduğunu görürüz. Dennett bilincin evrim sürecinde bir işlevi yerine getirmek için ortaya çıkmış olması gerektiğini düşünmektedir. Dennett, indirgemeci bir yaklaşımla bilinç problemini ele almış ve çoklu taslaklar modeli olarak bilinen bilinç modelini oluşturmuştur. Dennett modelini oluştururken *qualia*nın zihnin temel bir özelliği olmadığını belirterek birinci şahsa ait öznel deneyimlerin bir kullanıcı ara yüzü olduğunu, dolayısıyla *quali*anın özelliklerinin ve beynin *qualia*yı nasıl yarattığının açıklanmasının gerekli olmadığını savunur. Dennett bilinç modelinden *qualia*yı elemiştir ancak Dennett *qualia*nın varlığını yadsımaz, Dennett'a göre öznel deneyimlerimiz gerçektir ama düşündüğümüz şeyler değildirler. (Dennett, 2011, s: 35)

Dennett, çoklu taslaklar modelini oluştururken nörobilimin verilerini kullanır. Dennett'ın bilinç modelinde esas amaç beyin ve zihin düalizmini yaratan Kartezyen tiyatro yanılgısından kurtulmaktır. Dennett'a göre nörobilim dahi bilincin nöronal korelasyonlarını araştırarak Kartezyen tiyatro tuzağına düşmektedir. Kartezyen tiyatro beynin bilinci yarattıktan sonra beynin içinde bilinçli içeriklerin sunulduğu bir merkezin bulunduğu fikridir. Dennett bilinçli içeriğin beynin tek bir noktasında oluşturulmadığını vurgular. Çoklu taslaklar modelinde tüm beyne yayılmış aktivasyonlar sonucu bilinçli deneyimlerimiz ortaya çıkar. Bilinçli bir durumdayken duyularımız beynimizi her an inanılmaz miktarda bilgi bombardımanına tutar. Ancak beynimiz bu bilginin hepsini

işlemez, bilginin bir bölümü analiz edilir. Bu analizin içinde binlerce mikro an ve binlerce mikro yargı ve karar bulunur. Bu mikro yargılar ve kararlar bilinçli yargılar ve bilinçli kararlar değildirler: her yargı beynin konuyla alakalı uzmanlaşmış kısmında konumlanır. Mikro yargılar ve kararlar esnasında beynin daha önceki deneyimleri de kullanılır. Her sinyalde bir önceki sinyalin yargıları da işin içindedir. Beyinde paralel yürütülen aktiviteler sıralı oluşan aktiviteler değil, paralel bir şekilde aynı anda gerçekleşen aktivitelerdir. Bu çoklu yol süreçleri, tüm eklemeler, birleşmeler, düzeltmeler ve içeriğin üstüne yazmalar mili saniyelerde ve belirli bir sıra izlemeden gerçekleşerek sürecin sonunda zihin taslakları oluşturmaya başlar. Daha sonra bu taslaklar birbirleriyle rekabet ederek tüm beyinde kalıcı bir hakimiyet kurmaya çalışırlar. Bu hakimiyet bir sonraki adımda ne düşünüleceğini, neye dikkat edileceğini veya neye odaklanılacağını da belirler. Ancak bir taslağın ne zaman bilinçli olacağı bir muammadır. Dolayısıyla bilinç akışı tüm bu aktivitelerin geçmişe yönelik olarak beyinde iz bırakmasından ortaya çıkar. Belirli herhangi bir zaman sürecinde bilincinde olduğumuz taslak, sürece dair başlattığımız araştırmadan bağımsız olarak tamamlanmaz, devamlı bir revizyon süreci devam eder, dolayısıyla kurallara uygun kabul edilecek tek bir versiyon yoktur. Dennett, beyinde bir an önce olanın gelecekte ne olacağını da etkilediğinin açık olduğunu belirtir. (Dennett, 2011, s: 195-198 )

Ancak Dennett beynin bilince ev sahipliği yapmakla beraber bilincin saf beynin bir ürünü olmadığını da belirtir. Dennett'ın modelinde iki bileşenden oluşan bilincin ilk bileşeni beyin, ikinci bileşeni ise kültürdür. *Homo sapiens* binlerce yıl önce ilk ortaya çıktığında dünyanın biokütlesinin sadece binde birini oluştururken, bugün dünyanın biokütlesinin yüzde doksan sekizini oluşturmaktadır. Bu büyüme dünyanın tanık olduğu en hızlı gelişimdir. Dennett bu büyük değişimin sadece genlerdeki evrimle açıklanamayacağını ve bu gelişimin esas nedeninin kültür olduğunu belirtir. Dennett genlerle açıklanamayan bu gelişimi "mem"lerle açıklar. Memler kültür aktarım birimleridir. Dennett evrimin zeki tasarımlarının doğanın her noktasında bulunduğunu, insanın ise bu zekayı kültürel gelişimiyle idrak etmiş olduğunu belirtir. Dennett'a göre bilinç kültürle ve kültürü olanaklı kılmış dil ile oluşmuş bir kavramdır. (Dennett, 2017)

Dennett bilinç modelini, bilincin kültür bileşenini çoklu taslaklar modeline ekleyerek tamamlar. Dennett, kültürün beynin paralel donanımı üstüne kurulmuş bir sanal makine olduğunu iddia eder. Ancak kendi kontrol sistemlerine ve donanımı üstüne kurulu bir sanal makineye sahip olan bir şeyin bilinçli olabileceğini belirtir. Dolayısıyla Dennett bilinçli yapay sistemlerin teorik olarak olanaklı olduğunu düşünmektedir.

(Dennett, 2011, s: 328) Ancak Dennett teorik olarak olanaklı olsa dahi pratikte bilinçli yapay sistemlerin yaratılmasının önünde engeller olduğunu belirtir ve bilinçli yapay sistemler yaratmaya çalışmaktan daha çok işimize yarayacak zeki sistemler yaratmamız gerektiğini düşünür.

Bu çalışmanın ilk bölümünde yapay zekanın olanaklı olabileceğini savunan işlevselcilik incelenmiştir. Dennett yönelimsel duruşu araçsal bir yöntem olarak kullanarak yapay zekanın olanaklılığını epistemolojik bir probleme indirgemiştir. Yapay zeka çalışmaları da başlangıcından itibaren yapay zekanın olanaklılığına epistemolojik bir problem olarak yaklaşmıştır. İkinci bölümde yapay zeka çalışmalarının tanımlanan epistemolojik problemlere yaklaşım yöntemleri incelenmiştir ve yapay zekanın gelişimi, nörobilim araştırmacılarının beynin işlevlerini ve bilinci açıklamak için geliştirdikleri yöntem ve modeller ile paralel olarak değerlendirilmiştir. Üçüncü bölümde güçlü yapay zekanın olasılığına dair itirazlar günümüzdeki yapay zeka sistemleri üstünden incelenmiş, yine aynı bölümde öne sürülen itirazlara Dennett'ın karşı çıkıştığı noktalar açıklanmıştır. Dördüncü bölümde ise Dennett'ın zihin felsefesinin temelleri ve bilinç modeli açıklanmış ve Dennett'ın yapay zekanın olanaklılığını savunmasını Dennett'ın zihin felsefesiyle nasıl temellendirdiği araştırılmıştır. Dennett işlevsel bir bakış açısıyla insan bilicinin hem biyolojik yapısına dair bir model sunmakta, hem de bilincin insan kültürünün yaratılmasındaki işlevini açıklamaktadır. Ancak Dennett, işlevsel bir bakış açısına sahip olduğundan, *qualia*nın beyin tarafından nasıl ortaya çıkarıldığına bir açıklama getirmemekte, hatta *qualia*nın açıklanması gerekmeyen bir kavram olduğunu savunmaktadır. Ancak Dennett'ın da belirttiği gibi, öznel deneyimlerimiz, bir başka ifade ile *qualia*mız varoluşumuzun anlamını belirler. (Dennett, 2011, s: 35) Dolayısıyla *qualia*nın, söz konusu *qualia*yı oluşturan sistemle ilişkisini açıklamayan herhangi bir bilinç modeli birçok güncel soruyu cevaplamakta yetersiz kalmaktadır.

# 1. BÖLÜM

# YAPAY ZEKA VE İŞLEVSELCİLİK

## 1.1 YAPAY ZEKA KAVRAMININ ARKA PLANI

Yapay zeka kavramı Turing'in insanların yapabildiği hesapların bir makine tarafından da yapılabileceğini savunması ile doğmuştur. Turing insanların zihinsel yetilerinin makineler tarafından taklit edilebileceğini, insanlar gibi zeki davranışlar sergileyen makineler tasarlamanın mümkün olduğunu öne sürmüştür. İnsan zihninin yetilerinin makineler tarafından taklit edilebileceği düşüncesi ise zihnin çoklu gerçekleştirilebilir olduğu yani zihinsel yetilerin doğrudan organik bir beyin tarafından gerçekleştirilmesinin şart olmaması düşüncesini kabul eder. İşlevselcilik beyni zihinsel işlevleri gerçekleştiren bir organ, zihinsel yetileri ise işlevsel özellikler ve aynı zamanda çoklu gerçekleştirilebilir özellikler olarak tanımlayarak yapay zekanın temelindeki felsefi görüş olma niteliğini kazanmıştır.

1940'lı yıllara dek zihin konusunda Descartes'ın töz düalizmine dayanan ve zihnin fiziksel bir töz olmadığı savunan görüş hakim olmuştur. Bilişsel bilimler çerçevesinde Descartes'ın töz düalizmi ilk olarak Ryle tarafından eleştirilmiştir. Zihni makinedeki hayalet olarak adlandıran Ryle, zihnin bir varlık olduğu varsayımının bir kategori hatası olduğunu iddia etmiştir. (Ryle, 1949, s: 16) Ryle'a göre zihni bir varlık gibi düşünmek, zihni yanlış bir kategoride değerlendirmemize yol açmıştır. Ryle, zihinlerin beyinler de olmadığı vurgulamış, zihin sahibi varlıkların yaptıkları veya yapabileceklerine istinaden zihin durumlarına sahip olduğunu dolayısıyla zihin durumlarının davranışlar ve davranış yatkınlıkları ile analiz edilebilir olduğunu savunmuştur. Mantıkçı davranışçılık zihin durumlarıyla direkt olarak ilgilenmeyerek zihnin içsel niteliklerini bilişsel bilimlerden dışlamış, zihnin içsel niteliklerini davranışlara yani dışsal, gözlemlenebilir niteliklere indirgemiştir. Mantıkçı davranışçılığın karşılaştığı sorunlardan biri zihinsel durumların her zaman için belirlenmiş bir davranışsal çıktıya yol açmaması diğer bir sorunu ise zihin durumlarını diğer zihin durumlarına başvurmaksızın açıklayamaması olmuştur.

1950'li yıllarda bazı felsefeciler bilinçli süreçlerin beyin süreçleri olduğunu ve zihinsel durumların davranışların içsel nedenleri olduğunu savunmuş, zihinsel süreçleri beyin süreçleri ile özdeşleştirmişlerdir. Özellikle Place, yıldırımın elektrik yükünün hareketi

olduğunu, yıldırımın elektrik yükünün hareketinden bağımsız düşünülemeyeceğini aynı şekilde zihnin beynin süreçleri olduğunu ve zihnin de beyin süreçlerinden bağımsız düşünülemeyeceğini savunmuştur. (Place, 1956, s: 46) Place'a göre zihinsel bir durum beynin fiziksel kimyasal bir durumu ile özdeş olmalıdır. Özdeşlik kuramı olarak adlandırılan bu düşünce zihinlerin sinir sisteminin maddi özellikleri olduğunu ileri sürer. Başka bir ifadeyle özdeşlik kuramına göre belli bir zihinsel durumda olmak beynin belirli bir nöronal durumunda olmasından ibarettir. Acı çekmek beyindeki C-sinirlerinin ateşlenmesiyle özdeştir. Bu şekilde düşünüldüğünde zihinsel bir olayın maddi bir olaya yol açması, maddi bir olayın bir başka maddi olaya yol açması gibi açıklanabilir. Ancak Putnam özdeşlik kuramına karşı çıkarak sinir sistemi insanınkinden çok farklı olan ahtapotların acı çekmesinin düşünülebilir olduğunu, bu taktirde zihinsel durumların beynin fiziksel durumlarıyla özdeş olamayacağını belirtmiştir. (Putnam, 1967, s: 56) Putnam'a göre tüm zihinsel durumlar çoklu gerçekleştirilebilirdir. Bir başka ifadeyle zihinsel bir durum onu gerçekleştiren maddeyle özdeşleştirilemez ve zihinsel durumlar değişik düzenekler tarafından gerçekleştirilebilir.

Putnam'ın öne sürdüğü çoklu gerçekleştirilebilirlik argümanı işlevsel bakış açısında merkezi bir kavramdır. Aynı zamanda çoklu gerçekleştirilebilirlik argümanı yapay zeka çalışmalarının da temel kavramıdır, en nihayetinde yapay zeka kavramı belli zihinsel yetilerin makinelerde gerçekleştirilebileceği savına dayanmaktadır.

## 1.2 HESAPLANABİLİRLİK KURAMI

Günümüzdeki bilgisayar bilimlerinin temeli Turing'in 1936 yılındaki matematik alanında yayınlamış olduğu "Hesaplanabilir Sayılar. Karar Verme Probleminin Bir Uygulaması" adlı matematik makalesiyle atılmıştır. 1930'lu yıllarda matematikçilerin üstüne düşündüğü problemlerden biri matematiği mantık üstüne inşa etmeye çalışan Hilbert'in öne sürdüğü "karar verme problemi"ydi. Karar verme problemi, matematiksel bir önermenin verilmiş belli aksiyomlardan yola çıkarak doğru olup, olmadığına karar verebilen efektif bir prosedürün varlığının olasılığını sorgulayan teorik bir problemdi. Hilbert, matematiksel bir önermenin doğruluğunun aksiyomlarla göstermenin efektif bir prosedürü olduğu taktirde matematiği bir temele oturtmak için gerekli olan tüm aksiyomları tanımlamanın yeterli olacağına inanıyordu. (Penrose, 2015, s: 57)

Karar verme problemi, Gödel ve Church tarafından çalışılmış, Gödel ve Churh böyle bir prosedürün söz konusu olmadığını ve matematiksel önermelerin doğruluğunun sadece aksiyomlar üstüne kurulamayacağını göstermişlerdir. Aynı soruyla ilgilenen Turing ise probleme farklı bir açıdan yaklaşarak bir hesaplama yöntemi geliştirmiştir. Matematiksel olarak Gödel ve Church'le aynı sonuca ulaşan Turing'in makalesi özellikle kullandığı hesaplama yöntemi dolayısıyla bilgisayar bilimlerinin en temel makalelerinden birisi olmuştur.

Turing, makalesinde hesaplanabilir sayıların ondalık basamakları sonlu bir rakama dek gösterilebilen reel sayılar olduğunu belirtmiş ve hesaplanabilir sayılardan yola çıkarak hesaplanabilen matematiksel fonksiyonların neler olduğunu göstermiştir. Turing makalesinin önemi Turing'in hesaplanabilir sayıların hipotetik olarak otomatik bir makinede hesaplanabilir sayılar olduğunu göstermiş olmasıdır. Turing'in otomatik makinesi günümüzde “evrensel Turing makinesi” olarak bilinmektedir. Turing'in teorik olarak öne sürdüğü makine otomatik bir makinedir. Bu makine sonsuz uzunlukta hücrelerden oluşan bir banttan, bir sürücüden, bandın üstündeki sembolleri okuyup değiştirebilen sürücü kafasından ve istenilen hesabı yapan, belirlenmiş, sonlu kurallar dizisinden oluşur. Sonlu kurallar dizisi bandın üstündeki her sembol için tekrar tekrar kullanılabilir. Bandın üstündeki semboller belli sayıda ve önceden belirlenmiş sembollerdir ve bu sembollerin içinde başlangıç pozisyonunu, bitiş pozisyonunu ve bandın o hücresinin kullanılacağını belirten bir sembol bulunmak zorundadır. Başlangıç ve bitiş sembolleri arasındaki semboller girdiyi betimler. Sürücü hücrelerde sağa ve sola ilerleyebilmektedir. Sürücü kafası, sembollerin okuyabilme, değiştirebilme ve yeni sembolleri yazabilme özelliğine sahiptir. Turing, bir sayının hesaplanabilir olup, olmadığını gösterecek olan bu makinede, başlangıç, bitiş ve boşluk sembollerinin yanı sıra sadece bir ve sıfır sembollerini kullanmıştır. (Turing, 1936, s: 233-235)

Turing makinesinin çığır açıcı özelliği makinenin bir sembolü okuduktan sonra sembolü manipüle etme şeklidir. Turing makinesinde bir sembol okunduktan sonra sembole göre bir sonlu kurallar dizisi çalışır ve sonlu kurallar dizisi gerektiği taktirde sonsuz kez kullanılabilir. Bu noktada Turing'in motivasyonu matematikteki sonsuz sayıdaki teoremin ispatıyla tek tek uğraşmaktansa bu teoremleri belli kategorilere bölerek, her kategoriye uygulanabilecek bir otomatik kurallar dizisi yaratmak olmuştur. Turing makinesindeki kurallar dizisi günümüzde bilgisayarlarda kullanılan algoritmayı temsil eder. Turing, algoritmayı yaratırken aynı zamanda zor problemlerin nasıl basit

parçalara bölünebildiğini göstermiştir. Kısaca algoritma en genel haliyle özel problemlere uygulanabilen ve mekanik olarak o problemleri adım adım çözen bir reçetedir. Bu reçeteyi uygularken yaratıcılığa gerek yoktur. Kurallar bir insanın ya da makinenin uygulayabileceği türden kesinlik taşırlar. Sonlu kurallar dizisinin sonsuz kez kullanılabiliyor olması, Hilbert'in karar verme problemini durma problemine dönüştürmüştür. Eğer hesaplanmak istenen sayı hesaplanabilir bir sayı değilse hesap sonsuza dek devam eder. Turing makinesi işlemini tamamladığında, bandın üstündeki semboller cevabı temsil eder.

Turing makinesi hesaplanabilirlik konusunda bir referans noktası olmuştur. Günümüzde "Turing eşdeğer" olarak adlandırılan bu referans noktası, bir algoritma, herhangi bir donamında çalışabiliyorsa bu donamına karşılık gelen bir Turing makinesi olduğunu belirtir. Bir başka ifadeyle mekanik olarak yapılabilecek her hesaplama bir Turing makinesi tarafından yapılabilir. (Dennett, 2011, s: 252) Turing insanların yaptıkları hesapların hepsini hatta makinedeki sonsuz bandın, sonsuz bir hafızaya tekabül etmesinden dolayı daha fazlasını bir makinenin yapabileceğini düşünmüştür. İnsan zihninin yaptığı her türlü sembolik hesabın bir makine tarafından da yapılabileceği düşüncesi bilgisayar bilimlerinin doğması açısından bir dönüm noktasıdır. (Turing, 1950, s: 444)

İlk bilgisayarlar inşa edildikten sonra, Turing, 1948 yılında "Intelligent Machinery" makalesinde, evrensel makinenin öneminin açık olduğunu, evrensel makine sayesinde farklı işler için sonsuz sayıda makine gerekmeyeceğini, tek makinenin tüm işlere yeteceğini öne sürmüştür; "Farklı işler için farklı makineler üretmeyi ele alan mühendislik sorununun yerine, evrensel makineyi bu işleri yapmak üzere programlama işi geçmiştir." (Turing, 1948, s: 111) Dolayısıyla Turing'in, bilgisayarların yapabileceğini düşündüğü şeyler arasında insan zekasını taklit etmek de bulunmaktadır. (Nilsson, 2011, s: 265) Turing, 1950'de bir felsefe dergisinde yayınlamış olduğu "Computing Machinery and Intelligence" makalesinde insan zekasını bütünüyle mekanikleştirme olasılığını ayrıntılarıyla ele almış ve makinelerin insan zekasını taklit edebileceğini öne sürmüştür. Yine bu makalede, "makineler düşünebilir mi" sorusunun muğlaklığından dolayı, sorulması gereken doğru sorunun, "makineler insan zihnini taklit edebilir mi" olduğunu belirtmiş, "taklit oyunu" adı altında bir test tasarlamıştır. (Turing, 1950, s: 433) Turing testi olarak bilinen taklit oyunu, bir bilgisayarın bir insanı, dolayısıyla insanın düşünme sürecini, taklit edip edemeyeceğini sınayan bir testtir. Turing testinde sorgucu adı verilen bir insan, yazılı

mesajlaşmaya izin veren bir sistemle bir insana ve insan gibi düşünmek üzere tasarlanmış bir bilgisayara sorular sorar. Eğer sorgucu, verilen cevapların hangisinin insana, hangisinin bilgisayara ait olduğuna karar veremezse, bilgisayar testi geçer. Bilgisayarın testi geçmesi ise testi geçen bilgisayarın insanın bilişsel özelliklerine denk bilişsel bir sisteme sahip olduğunu gösterir. Turing testinde önemli olan, girdi ve çıktılar arasındaki bağlantıdır ve bilgisayarın iç temsillerinin ne olduğu önemsizdir. Turing'in 1950 yılında yayımlanan makalesindeki yapay zeka çalışmalarına yön veren bir başka önemli nokta, insan düzeyinde düşünsel beceriler taşıyan programları üretmeye nasıl başlanabileceği konusunda sunduğu bir öneridir ki bu öneri yetişkin insan zihnini taklit eden bir program yaratmaya çalışmak yerine çocuk zihnini taklit eden, böylelikle eğitimden geçirilebilecek programlar yaratmaya çalışmaktır. (Turing, 1950, s: 460)

## 1.3 İŞLEVSELCİLİK

Turing, Turing makinesinde bir program işletilirken makinenin girdiği durumları girdi-çıktı ilişkileri ve makinenin girdiği diğer durumları kullanarak tanımlamış olması zihin felsefesinde işlevsel yaklaşımlara bir örnek teşkil etmiştir. İşlevselcilik zihinsel kavramları beyin ile özdeşleştirmekten daha çok Turing makinesinde olduğu gibi, zihinsel durumların beyindeki nöronal durumların işlevleriyle belirlenebileceğini öne sürer. Nasıl ki kalbin ne olduğunu belirleyen kan pompalama özelliği yani işleviyse ve bu işlev kalbin doğru nedensel rolü oynamasıyla kazanılıyorsa zihinsel özelliklerde işlevsel özelliklerdir.

### 1.3.1 Makine İşlevselliği ve Nedensel İşlevselcilik

Putnam tarafından ortaya atılmış olan makine işlevselliği doğrudan Turing makinelerinden ilham alarak insan zihninin olasılıklı geçişlere sahip bir otomat olduğunu savunmuştur. Turing makineleri deterministtik sistemlerdir, bir Turing makinesinin bir durumdan diğer duruma geçiş olasılığı sadece sıfır ve birdir ancak olasılıklı bir Turing makinesinde geçiş olasılıkları sıfır ve bir arasında değerler alır, dolayısıyla Turing makineleri olasılıklı otomatların özel bir durumudur. (Putnam, 1967, s: 54) Makine işlevselliği zihinsel durumları ortaya çıkaran alt durumları sonlu bir durumlar dizisi olarak tanımlar. Bir organizmanın girdileri duyusal algılar, çıktıları ise

motor hareketlerdir. Putnam girdi ve çıktı arasındaki zihinsel durumların ise aralarında olasılıklı geçişlere sahip otomat durumları olduğunu savunmuştur. Bu durumlarsa, Turing makinesinde olduğu gibi bir makine tablosuyla tanımlanabilir. Makine tablosundaki durumlar sadece girdi ve çıktılar ile bağlantılı değildir, hatta bu durumlar girdi ve çıktılarla bağlantılı olmayan ve sadece diğer durumlarla bağlantılı olan durumlar olabilir. Makine bir çıktı üretmiyor olsa bile makinede etkin durumların olması olasıdır. İçsel temsiller, bir başka ifadeyle zihinsel durumlar ise girdi ve çıktı arasındaki makine tablosundaki etkin durumların toplamlarının yarattığı ifadelerdir. Putnam girdi ve çıktılar arasındaki gerçekleştiricilerin ne olduğunu bilmediğimizi varsayar, acı durumuna yol açan girdiler ve acı durumunun çıktılarının acı ile bağlantılarını bilebilir ancak yine de acının ne olduğunu bilmeyebiliriz. (Putnam, 1967, s: 57)

Putnam bir organizmanın sistemin bütünü olarak ele alınması gerektiğini de öne sürmüştür, örneğin acı içinde olmak organizmanın bütününün bir işlevsel halidir. Putnam'ın makine işlevselliğinde acı hissetmeye meyilli tüm organizmalar olasılıklı otomatlardır. Acı hisseden tüm organizmalar ise çevreleriyle nedensel bir etkileşim içindedir. Dolayısıyla acı organizmanın çevresiyle etkileşiminde nedensel bir rol oynamaktadır. Acı hissetme yetisi olan her organizma ise en az birer zihinsel temsile sahiptir, acı işlevsel bir rol oynamaktaysa acının işlevsel rolünü organizma için harekete geçirecek bir temsile ihtiyaç olacaktır. Putnam temsilleri makinenin etkin içsel durumlarının bir toplamı olarak tanımlamıştır ve girdi ile çıktı arasındaki durumlar temsilleri oluşturur. Bir temsil olduğunda organizma içine girebileceği olasılıklı durumlara sahip olur. Dolayısıyla acı bu durumlar üstünden saptanabilir bir durumdur. Bu noktada Putnam organizmanın bir bütün olduğunu, organizmanın bir alt kümesinin bir organizma olmadığının altını çizmiştir. (Putnam, 1967, s: 54) Sonuç olarak acı organizmanın işlevsel bir hali olduğunda acı sadece beyindeki nöronal durumlarla özdeştirilemez; acı ancak girdiler, çıktılar ve organizmanın işlevsel hallerinin ilişkileriyle tanımlanır. Putnam ayrıca bir organizmanın belli işlevsel temsillere sahip olmasının ancak bu organizmanın tercih yapabilmesiyle anlam kazanacağını belirtir, tercih yapabilmek ise deneyimden öğrenmeyle alakalıdır.

Bir başka işlevsel kuram ise nedensel işlevselciliktir. Nedensel işlevselliğe göre zihinsel özellikler işlevsel özelliklerdir ve zihinsel durumlar diğer zihinsel durumların, algıların ve davranışların nedensel ilişkileriyle belirlenir. (Levin, 2021, s: 2) Mantıksal davranışçılar acıya neden olan bir yaralanmanın çığlık atmak gibi davranışsal tepkilere yol açacağını savunurken Lewis, acıyı nedensel bir rol oynayan bir durum

olarak tanımlamıştır. (Lewis, 1980, s: 217) Acının yol açtığı zihinsel durum diğer zihinsel durumları da ortaya çıkaracak, davranışlar ise zihinsel durumların sonucu olacaktır. Örneğin acı inancının zihinsel durumu çığlık atmanın ötesinde acıyı dindirmenin yollarına dair inançları içeren zihinsel durumları tetikleyecek, yaranın tedavi edilmesi gerektiği zihinsel durumunun ortaya çıkmasına neden olacaktır. Dolayısıyla yaraya yara bandı yapıştırmak bu zihinsel durumun sonucu olan bir davranış olarak ortaya çıkacaktır.

Lewis, "Delinin acısı ve Marslının acısı" adlı makalesinde bir delinin acıya vereceği davranışsal tepkilerin normal bir insanınki gibi olmayacağını ancak bu durumun delinin davranış çıktılarına özel istisnai bir durum olduğunu, delinin yanlış nedensellikler sergilediğini belirtir. Sinir sistemi hidrolik mekanizmalardan oluşan bir Marslının ağrısının ise insan sinir sistemindeki C-sinir liflerinin gerçekleştirdiği ağrıdan çok farklı bir mekanizma üstünden gerçekleştirileceğini, Marslının davranış çıktılarının farklı olabileceğini ancak acının Marslı vücudunda da aynı nedensel rolleri oynadığından dolayı Marslıya insanınki gibi bir his vereceğini savunmuştur. Verdiği örnekler üstünden Lewis, özdeşlik kuramının deli adamın acısını açıklayabildiğini ancak Marslı ağrısını açıklayamadığını, işlevselliğin ise Marslı acısını açıklayabildiğini ancak delinin acısını açıklayamadığını belirtmiştir. Ancak delinin acısını açıklanamaması işlevsellik açısından bir sorun oluşturmamaktadır çünkü işlevselcilik nedensellik üstüne kurulmuştur ve delinin normal insanlar için belirlenmiş nedensel ilkelere uymadığı açıktır. Bu noktada Lewis, Putnam'ın çoklu gerçekleştirilebilirlik kavramının özdeşlik ilkesini yanlışlamadığını, insana ait zihinsel bir durumun insan beyniyle bir durumla özdeş, Marslı zihinsel durumunun ise Marslının hidrolik sinir sisteminde bir durumla özdeş olacağını dolayısıyla önemli olanın sunulan önermelerde alanın belirlenmesi olduğunu belirtmiştir. (Lewis, 1980, s: 220)

Lewis zihinsel terimlerimizin zihinsel durumların nedensel ilişkileri, duyusal algılar ve davranış çıktılarını halk psikolojisinin terimleriyle açıkladığımızı vurgulamıştır. (Lewis, 1972, s: 49) Duyusal algı ve davranış çıktıları arasındaki tüm zihinsel durumların aynı anda analiz edildiği durumda ise halk psikolojisinin kavramlarının analizden elenebileceğini ve böylece duyusal algı ve motor çıktı arasındaki tüm durumların nöronal etkinliklere indirgenebileceğini belirtmiştir. Örneğin vücuttaki bir yaralanma acıya yol açar, acı hissi vücutta bir şeylerin ters gittiği inancını ve acıyı dindirme isteğini oluşturur, acıyı dindirme isteği ise yaranın tedavi edildiğinde acının dineceği inancıyla etkileşime girer ve yaralanan kişi yarasına yara bandı yapıştırır. Bu cümlede

acıyı bir x durumu olarak, diğer inanç ve zihin durumlarını da nedensel durumlar, örneğin y, z ve t durumları olarak tanımlarsak ve cümlesindeki tüm zihinsel durumları bir bütün olarak yazarsak, cümle ∃x∃y∃z∃w formunda olacaktır. Bir Ramsey cümlesi olan bu cümlede zihin durumları sadece nedensel ilişkileri gösterilerek yazılmış, böylece cümleden zihinsel durumların halk dilindeki tanımlamaları çıkarılmıştır. Dolayısıyla zihinsel durumlar uyaran ile davranış çıktısı arasındaki nicelleştirmelere indirgenmiştir. (Levin, 2021, s: 21) Lewis'in işlevselliği analitik işlevselcilik olarakta tanımlanmaktadır.

### 1.3.2 Rol İşlevselliği

Analitik işlevsellik ve makine işlevselliği zihinsel durumları gerçekleştiricileri üzerinden tanımlamaktadır. Nedensel ilkelere uyan gerçekleştiriciler ile zihinsel özellikler ortaya çıkar. Rol işlevselliği ise zihinsel kavramları oynadıkları roller üstünden tanımlamayı tercih eder. Rol işlevselliğinin gerçekleştirici işlevselliğinden ayrıldığı nokta gerçekleştirici düzeyi ile zihinsel kavramların düzeyini ayırt etmesidir. Rol işlevselliği, acının kendisine has özelliklerini sorgulayarak acıyı, acı rolünü üstlenen üst seviye bir durum olarak tanımlar. (Levin, 2021, s: 17)

Rol işlevselliğinin savunucularından Fodor, temsili zihin kuramında zihinsel durumların üst düzey durumlar olduğunu belirtir. Halk psikolojisinin sezgi, inanç, istek, arzu gibi kavramlar üstünden bir model oluşturduğunu belirten Fodor, işlevsel bir zihin modelinin üst seviyedeki özellikler üstünden tasarlanması gerektiğini savunmuştur. İnanç, istek ve arzular gibi zihinsel olayları önermesel tutumlar olarak tanımlayan Fodor, önermelerin doğruluk değeri taşıyan üst seviye soyutlamalar olduğunu belirtir. Fodor önermesel tutumların bir davranışı açıklamak veya bir davranış öngörüsünde bulunmak için kullanıldığını ve önermesel tutumların öngörülerde kullanımının doğru sonuçlar verdiğini iddia etmiştir. (Fodor, 1987, s: 6) Fodor'a göre önermesel tutumlar özneler ile zihinsel temsiller arasındaki bağlantılardır.

Fodor'un modelinde zihinsel temsiller merkezi bir konumdadır. Zihinsel temsiller semantik özelliklere sahip zihin nesneleridir. Herhangi bir "p"ye inanmak Fodor'a göre p'nin semantik özelliklerini taşıyan zihinsel temsil ile inanma ilişkisi kurmak anlamına gelir. Fodor zihinsel temsillerin örnekler üstünden oluşturulan tipler olduğunu öne sürmüştür. Temsiller kalıcı durumlar veya olaylar üstünden tanımlanan zihinsel etkinlikler üstünden örneklenmektedir. Fodor'a göre düşünmek zihinsel etkinlikler

zincirlerinin zihinsel temsiller üstünden örneklenmesidir. Fodor'un temsili zihin kuramında düşünce alanını belirleyen nesneler zihinsel temsiller, düşünmek ise bu nesnelere uygulanan önermesel tutumlardır. Fodor, zihnin semantik anlamları olan sembolleri manipüle eden bir Turing makinesine eşdeğer olduğunu iddia etmektedir. Turing makinesinin algoritmasını oluşturan kurallar Fodor'a göre beynin içsel yapısına gömülü kurallardır. (Fodor, 1994, s: 7-9) Turing makinesindeki semboller yorumlanmamış sembollerdir, bir başka ifadeyle semboller anlamdan bağımsız manipüle edilmektedir. Zihin bir hesaplama yaptığı ve bunu yorumlanmamış semboller üstünden yaptığı düşüncesi aynı zamanda zihin temsillerinin birbirileriyle belli ilişkiler içinde kullanıldığı anlamına gelmektedir. Fodor'a göre bahsi geçen bu ilişkiler insan dilinin temel özelliklerinde tanımlanmıştır. Fodor'un düşüncenin dili teorisi tüm insanların zihninde ortak ve kurallı bir yapı olduğunu ve söz konusu yapının konuşulan dilin yapısıyla aynı yapıya sahip olması gerektiğini belirtmiştir. Her ne kadar konuşulan dil düşüncenin dilinin yapısıyla aynıysa da düşüncenin dilinin konuşulan dilden daha geniş bir kapsamı vardır.

Fodor'un temsili zihin kuramında zihnin işleyişi alt seviye düzeydeki semantik anlam taşıyan temsillerin bir sembol gibi kullanılarak düşüncenin dilindeki cümleler üzerinden gerçekleşen hesaplamalar ile ifade edilebilir. Fodor bu hesaplamaların algoritmasının ise psikoloji bilimi tarafından araştırılmış olduğunu ve psikolojinin bu algoritma hakkında doğru öngörüleri olduğunu savunmuştur. Önermesel tutumları içeren psikoloji ise beynin yapısından soyutlanmış üst düzey bir bilimdir. Fodor, zihnin işleyişini program düzeyinde tanımlamıştır, bir başka deyişle zihnin bir Turing makinesindeki gibi kurallardan oluşan bir yazılım, beynin ise bu yazılımın içine gömüldüğü fiziksel yapı olduğunu iddia etmiştir. Bir başka ifadeyle Fodor, zihin kuramını inanç, istek arzu gibi halk psikolojisi kavramları üstünden kurmaya çalışmış ve temelinde nedensel ilkelerle çalışan beyin olduğu sürece halk psikolojisinin davranışları açıklamak ve öngörmekte ki başarısına vurgu yaparak halk psikolojisinin terimlerinin bir zihin kuramının temel kavramları olabileceğini iddia etmiştir.

### 1.3.3 Dennett'ın Üç Duruşu ve İşlevselcilik

Dennett, arzu, inanç gibi zihinsel durumların anlamlarını yönelimsel duruş kavramı altında incelemiştir. Dennett'a göre bu kavramları atfederek birçok organizma ya da yapay sistemin davranışlarını bu kavramlar üstünden öngörebiliriz ancak sistemlere

yönelimsel duruşu atfeden bizizdir. Bir sistemin davranışları yönelimsel duruşla açıklandığında ya da yönelimsel duruşla öngörüldüğünde bu sistem yönelimsel duruşu olan bir sistemdir. (Dennett, 1987, s: 2-3) Dennett yönelimsel duruşun yanı sıra fiziksel duruş ve tasarım duruşu olarak iki duruş daha tanımlar. Fiziksel duruş bir sistemin hareketini fizik yasaları üstünden anlamaya ve öngörmeye çalışır. Tasarımsal duruş ise bir sistemin tasarımındaki belli bir amacı yerine getirmek üzere tasarlanmış olduğunu düşünerek sistemin hareket veya davranışları anlayıp, öngörür. (Dennett, 1987, s: 5)

Dennett'a göre bir sisteme hangi duruşla yaklaşacağımızı amaçlarımız doğrultusunda biz seçeriz. Elimizdeki bir taşı ve çalar saati bıraktığımızda bu iki nesnenin yere düşerken nasıl bir yol izlediğini öngörmenin en kolay yöntemi fiziksel duruşu benimsemektir. Ancak çalar saatin alarmının çalmasını açıklamak için tasarımsal duruşu seçmek daha iyi bir yoldur. Çalar saat bir amaca göre tasarlanmıştır ve çalar saatin ne zaman çalacağını öngörmek için fiziksel bir duruştansa tasarımsal duruş tercih edilir. (Dennett, 1987, s: 4) Aynı şekilde eğer satranç oynayan bir bilgisayarın davranışlarını ön görmek istersek tasarımsal duruşla satranç algoritmasını inceleyebiliriz ama eğer satranç oynayan programa yönelimsel bir inanç atfedersek, yani bilgisayarın oyunu kazanma arzusunda inançları doğrultusunda hareket ettiğini varsayarsak satranç oynayan bilgisayarın hareketlerini daha iyi açıklayabiliriz ki bu bilgisayarın davranışlarını öngörmekte yönelimsel duruşun, tasarımsal duruştan veya fiziksel duruştan daha iyi bir seçim olduğunu açıkça ortaya koyar. (Dennett, 1987, s: 5) Dennett, aynen bilgisayara yönelimsel duruş atfettiğimiz gibi hayvanlara ve insanlara da yönelimsel duruş atfetmenin onların davranışlarını açıklamak için iyi bir yöntem olduğunu savunmuştur. Yönelimsel duruş atfederek sistemlerin halk psikolojisi kavramlarıyla açıklanması Dennett'a göre kullanışlı bir metot olmasının ötesinde halk psikolojisinin uygulandığı sistemlerdeki düzenlilikler o sistemlerin davranışlarını doğru öngörmemizi sağlamaktadır. Dennett, bir sistemdeki düzenliliklerin o sistemin davranışı öngörmekteki esas nokta olduğunu savunur. Bir sisteme yönelimsel duruş atfetmek için o sistemin kendi çevresinde belli inanç ve arzular doğrultusunda hareket eden bir ajan olduğu varsayılır. Ancak Dennett, bir sistemin gerçekten yönelimsel bir duruş sergilemesiyle bir sistemin davranışının yönelimsel bir duruş atfedilerek öngörülmesi arasında fark olduğunu da belirtir. İnsanların yönelimsel duruşa sahip olduğunu ve insanların rasyonel ajanlar olduğunu varsayarız. Aynı şekilde hayvanların davranışlarını öngörmek için onlara yönelimsel bir duruş atfederiz. Dennett'a göre belli sistemlere belli özellikler atfedilmesini belli bir

süreklilik kabul ederek açıklayacak teorik bir prensip olmadığı gibi yönelimsel duruşa sahip olan sistemler ile yönelimsel duruşa sahipmiş gibi davranan sistemleri ayırt edecek teorik bir prensip de yoktur. Bir başka ifadeyle inanç ve arzularla tanımlanabiliyor olmak her zaman inanç ve arzulara sahip olmak değildir ve bir sistemin gerçekten inanç ve arzulara sahip olup olmadığını anlamamızı sağlayacak bilimsel bir yöntem yoktur. (Dennett, 1987, s: 8)

Dennett, birinci dereceden yönelimsel duruşa sahip olmayı sistemin davranışının inanç ve arzular üstünden öngörülebilir olmasıyla tanımlar. İkinci dereceden yönelimsel bir duruşa sahip olmaksa sistemin kendi inanç ve arzularına dair inançları olmasıyla tanımlanır. Dennett gerçek anlamda yönelimsel duruşa sahip olan sistemlerin ikinci dereceden yönelimsel duruşa sahip olan sistemler olduğunu belirtir. Bu şekilde Dennett insan zihni ile yaratık zihni arasında bir ayrım yaparak yaratık zihninin yaratığı yönlendiren temsillerinin açıklaması için tasarımsal duruşun yeterli olduğunu savunur. Bir varlığa yönelimsel duruş atfetmek Dennett'a göre o varlığın davranışını açıklamaya veya öngörmeye yönelik araçsal bir tutumdur. Dennett inanç ve arzuların herhangi bir varlığa atfedilebileceğini ancak insan zihni ile diğer yaratıkların zihni arasındaki süreksizliğin inanç ve arzuların sadece insan zihninin içerikleri olduğunun göstergesi olduğunu düşünmektedir.

Dolayısıyla herhangi bir sistemin davranışlarını açıklamak veya sistemin davranışlarını öngörmek için o sistemi en kolay hangi duruşla inceleyebiliyorsak sisteme o duruşu atfedebiliriz. Bu sistemin yapay veya doğal bir sistem olması ya da eğer sisteme yönelimsel duruş atfedeceksek sistemin gerçekten inanç veya arzulara sahip olması önemli değildir. Bir sisteme bir duruş atfetmekte ki amaç söz konusu sistemin davranışlarını açıklamak ve o sistemin hareketlerini öngörmektir. Bir sistemin hareketlerini öngörmek ise sistemin belli düzenliliklere sahip olmasını gerektirir. Düzenliliklere sahip bir örüntü taşıdığı bilgi kaybolmaksızın daha küçük bir örüntüye dönüştürülebilir. Dennett önemli olanın sistemin hareketinin öngörülmesi ve açıklanması olduğunu, dolayısıyla bizim sisteme atfettiğimiz duruşun kavramlarının ontolojik özelliklerinin önemsiz olduğunu düşünmektedir. En nihayetinde bir sistemin davranışlarını öngörmek epistemolojik bir problemdir. Bir başka ifadeyle inanç, arzu gibi zihinsel olaylar bir sistem için tanımlanması mümkün olan ve epistemolojik açıdan faydalı kavramlardır. (Dennett, 1987, s: 8)

Bilimde belli seviyelerde fenomenler farklı düzenlilikler gösterir. Her seviye de fenomenler farklı düzenliliklere sahip olabilir ve bilimler kendi alanları olan düzeylerde

o düzeyi betimleyen kavramlarla iş görürler. Bir sistemde karmaşıklık arttıkça o sistemdeki fenomenlerin kendine has özellikler sergilemeye başlaması ise belirme (emergence) kavramı ile açıklanır. (Connor, 2020, s: 1) Örneğin sudaki yüzey gerilimi, akışkanlık gibi özellikleri su moleküllerinin yapısıyla ve temel kimya biliminin kavramlarıyla açıklanamaz. Suyun özellikleri birçok su molekülünün birbiriyle karışık etkileşimleriyle belirir ve bu yeni özellikleri açıklamak için temel kimya biliminde olmayan kavramlar ve düzenlilikler kullanılır. Suyun özellikleri mikro düzeyde su moleküllerinin sahip olmadığı fakat makro düzeyde birçok su molekülünün arasındaki etkileşimle beliren yeni özelliklerdir. Mikro düzeyde akışkanlık veya yüzey gerilimi kavramları herhangi bir anlam taşımazken makro düzeyde beliren yeni özellikler suyun davranışını açıklamak için gereklidir. Akışkanlık kavramı alt düzeyde su moleküllerin yapısı incelenerek tahmin edilmez bir özellik olsa dahi üst düzeyde beliren akışkanlık alt düzey özelliklerden, birçok su molekülünün birbiriyle etkileşimiyle açıklanabilir. Bilimde düzeyler fenomenlerin farklı özelliklerine göre fenomenleri farklı kavramlarla ve düzenliliklerle açıklar. Üst düzey düzenlilikler genelde alt düzey düzenliliklere doğrudan indirgenemez. Bütünleştirici düzeyler (Integrated levels) daha düşük seviyenin önceden var olan fenomenlerinden beliren yeni fenomenleri içerir.

Dennett zihnin yapısının anlaşılması için beynin birden fazla düzeyden oluştuğunu düşünmemiz gerektiğini ve her düzeyin kendine has düzenlilikleri olduğunu savunur. Düzeyler arasında düzenliliklerin farklı düzenlilikler olarak gözlemlenmesi her düzeyde farklı kuralların geçerli olduğu anlamına gelir. Bir başka ifade ile beyindeki her farklı düzeyde beynin karmaşıklık derecesinden dolayı farklı düzenlilikler ile açıklanan farklı özellikler belirmektedir. Dolayısıyla belirli bir düzeyde hangi duruş daha pratikse söz konusu düzeyde o duruşu benimseyebiliriz. Örneğin nöronal seviyede fiziksel duruşu, belirli bir işlevi yerine getiren bir nöron kümesi için tasarımsal duruşu ve üst seviyelerde organizmanın davranışlarını açıklarken ise yönelimsel duruşu benimseyebiliriz. (Dennett, 2008, s: 191) Zihni anlamaya çalışırken zihne fiziksel duruşla yaklaşmak muhtemelen bize fazla bilgi vermeyecektir. Ancak yönelimsel duruş ile zihnin özelliklerini açıklarken de fiziksel duruş ile açıklayamadığımız özelliklerle karşılaşırız ve bu özelliklerin beynin karmaşıklığı sonucu ortaya çıkan yeni özellikler olduğunu düşünürüz. Bu özelliklerin, örneğin bilincin, ontolojik olarak fiziksel olmayan bir özellik olması da bazı filozoflarca kabul edilen bir düşüncedir.

Belirme zayıf belirme ve güçlü belirme olmak üzere ikiye ayrılır. Zayıf belirmede ortaya çıkan, beliren yeni özelliğin açıklaması alt düzeyde iş gören kavramlar türünden hesaplanamazsa bile açıklanabilir; bir başka ifade ile üst düzeydeki fenomenlerinin özellikleri alt düzey özelliklere indirgenebilir. Üst düzey özelliklerin alt düzey özelliklerle açıklanabilir olması bu özelliklerin alt özelliklere üst-bağlanır olarak tanımlanmasını sağlar. (Kim, 1984, s: 48) Üst düzeyde yeni özelliklerin belirmesi genel olarak çok karmaşık ve içsel etkileşimleri lineer olmayan sistemlerde gözlemlenir. Bu tip sistemlerde beliren yeni özellik alt düzeyin işleyişini tanımlayan yasalarla öngörülemeyen bir özellik olabildiği gibi hesaplama karmaşıklığından dolayı üst düzey özelliğin düzenliliği alt düzey özelliklerden hesaplanamayabilir ancak yine de teorik olarak beliren yeni özellik alt düzeydeki özellikler ve düzenlilikler cinsinden açıklanabilir. Güçlü belirmede ise yeni özelliğin alt düzey kavramlar ile açıklanması mümkün olmamakta, üst düzeyde beliren yeni özellik alt düzey bilimin kavramlarına indirgenememektedir. Temel fizikte kuantum dolaşıklığı (*quantum entanglement*) bir çeşit güçlü belirmedir. Güçlü belirme bir sistemin alt düzey bilim kavramlarınca açıklanamayan ontolojik özelliklerinin belirmesi olarak ta yorumlanabilir. (Connor, 2020, s: 14, 23)

Dennett zihinsel özelliklerin zayıf beliren özellikler olduğunu dolayısıyla zihinsel özelliklerin hesaplanamaz olsa da teorik olarak alt seviye nöronların özellikleriyle açıklanabileceğini savunmaktadır. Dennett'ın bu tezi yaşam oyunu inceleyerek daha iyi anlaşılabilir. Yaşam oyununu ızgaralar üstündeki damalar ile başlar. Her damanın yerel çevresiyle etkileşimini belirleyen dört basit kural vardır ve bu kurallar lineer olmayan etkileşimler içerir. Oyunun ilk aşamalarında ortaya çıkan düzenlilikler fiziksel duruşla hesaplanabilir özelliklerdir. Oyun ilerledikçe oyundaki yapılar yeni düzenlilikler sergilemeye başlar. Örneğin oyun belli bir karmaşıklık seviyesine ulaştığında yanıp sönen düzenlilikler ya da belli bir düzende hareket eden dama grupları gözlemlenebilir. Sistem karmaşıklaştıkça ortaya çıkan bu tip yeni düzenlilikler teorik olarak oyunun başlangıç durumları ve dört kural ile açıklanabilir olsa da belli bir noktadan sonra özellikle başlangıç koşullarına erişim imkanı kalmadığından ortaya çıkan düzenlilikler alt düzey düzenlilikler ve dört kural aracılığıyla açıklanamaz hale gelir. Belli bir karmaşıklık derecesindeki bu yeni düzenlilikleri belirmiş özellikler olarak tanımlarız. Karmaşıklık seviyesi iyice arttığında, özellikle tüm sistemin düzenliliğinden daha çok sistemin içinde ortaya çıkmış bir düzenlilikten bahsediyorsak bu düzenliliği yönelimsel duruşa geçerek inceleyebiliriz. (Dennett, 2008, s: 200) Eğer incelediğimiz düzenlilik oyunun içindeki bir grup damanın oluşturduğu bir düzenlilik ise söz konusu

damaların oluşturduğu sistem çevreyle etkileşimlerinden dolayı sanki yukarıdan aşağı bir nedensellik sergiliyormuş gibi gözükür. Örneğin kendi kendini organize eden biyolojik bir yapı hücrelerden oluşmaktadır ancak çevre koşullarına karşı koyması ancak bütünün hücreleri düzenleyecek şekilde bir sistem oluşturmasıyla mümkündür. Bir başka ifadeyle seçmiş olduğumuz biyolojik sistem sadece hücrelerin alt düzey özellikleri ve düzenlilikleriyle değil aynı zamanda bütünün, çevreye etkileşimi dolayı oluşan, üst düzey özelliklerin alt düzey özelliklerine etkileriyle tanımlanabilir. Bütünün üst düzey özelliklerinin alt parçaları etkilemesi aşağı nedensellik olarak tanımlanır. (Connor, 2020, s: 14) Üst düzeydeki sistem ile bu sistemin alt düzey yapıları arasında döngüsel bir etkileşim oluşur. Dolayısıyla bu sistem nedensellik ilkesini ihlal ediyormuş gibi yorumlanabilir. Biyoloji çevrenin etkileri ile organizmaların kendi kendini organize etmesini incelerken yukarıdan aşağı nedenselliği varmış gibi kabul eder. Bu durumda yaşam oyunundaki zayıf belirlenimin beş önermesinin ( 1 – Üst düzey özellik ve düzenlilikler alt düzey özellik ve düzenliliklere üst bağlanır. 2 – Beliren özellik ve düzenlilikler gerçek özellik ve düzenliliklerdir. 3 – Beliren özellikler nedensel etkiye sahiptir. 4 – Beliren özellik ve düzenlilikler alt düzey özellik ve düzenliliklerden farklıdır. 5 – Alt düzey düzenlilikler nedensellik yasalarına uyar. ) sonucu fazla-belirleme (overdetermination) olacaktır. (Kim, 1998, s: 44) Dennett'a göre fazla-belirme problemi tüm sistem fiziksel duruşla incelendiğinde buharlaşan bir problemdir. Ancak yaşam oyunu lineer olmayan kuralları ve karmaşıklığı ile ve karmaşıklık arttıkça oyunda yeni düzenliliklerin ortaya çıkmasından dolayı, oyunun başındaki ilk durumlara da sistemin o anki durumundan başlayarak ulaşma imkanı olmadığından dolayı pratikte fiziksel duruşla incelenmesi mümkün olmayan bir sistemdir. Bu durumda araçsal olarak duruşumuzu değiştirip önce tasarımsal duruşla oyunda ortaya çıkan bir sistemin yeni düzenlilikleri incelememiz mümkündür. Oyunda insan seviyesinde karmaşık sistemler oluştuğunda ise duruşumuzu yönelimsel duruş olarak tayin etmemiz sistemdeki yeni düzenlilikleri keşfetmemiz için gereklidir. Yaşam oyununda fiziksel duruşun ötesinde bir duruş seçip, tüm ekrandan daha çok ekranın bir bölümündeki düzenlilikleri araştırmaya başladığımızda yeni düzenliliklerin güçlü belirimle ortaya çıkan düzenlilikler olduğunu düşünebiliriz ancak bu düzenlilikler tamamen bizim seçtiğimiz duruşla alakalıdır. (Dennett, 2008, s: 195-199)

Dennett zihnin beynin karmaşıklığından dolayı beliren bir özellik olduğunu ancak belirmenin zayıf bir belirme olduğunu düşünmektedir. İndirgemeci bir filozof olan Dennett, zihnin özelliklerinin beynin yapısı ve en temel seviyede nöronlarla bağlantılı olduğunu ve beynin çok düzeyli bir yapı olarak farklı düzeylerde farklı düzenlilikler

sergilediğini savunur. Alt seviyede nöronlar olmak üzere beyin hiyerarşik bir düzende gitgide karmaşıklaşan seviyelerden oluşmaktadır. Zihin ise organizasyonun en üst seviyesinde beliren bir özelliktir. Dennett karmaşık bir sistem olan beynin değişik seviyelerde farklı düzenlilikler sergilemesinin olası olduğunu düşünür. En üst düzey yönelimsel duruşla incelediğimiz zihinsel düzeydir. (Dennett, 2008, s: 195) Bir alt seviyede kendi işlerini yapan ve kendi içinde özerk ve kendi düzenlilikleriyle işleyen modüller daha alt seviyede nöronal hesapları yapan kolonlar ve en alt seviyede ise tek başına özerk hücreler olan nöronlar bulunur. Karmaşıklığın belli seviyelerinde ortaya çıkan düzenlilikleri anlamak zihnin indirgemeci bir açıklamasını yapmak için gereklidir. Dennett yönelimsel duruş atfedilen en üst seviyenin altındaki alt-insan seviyelerinin özellikle mekanistik olduğunu da belirtmiştir. (Dennett, 2007, 78-79) Dolayısıyla Dennett zihni araştırırken yönelimsel duruşu epistemolojik sebeplerle tercih eder, bilimin de nihai amacı sistemlerin epistemolojik özelliklerini açıklamaktır ve Dennett duruşunu tanımlayarak kullandığı zihinsel özelliklere ait kavramların bir araç olduğunu belirterek zihnin ontolojik incelemesini konunun dışında bırakmaktadır.

### 1.3.4 Homonkular İşlevselliği

Dennett'ın zihne düzeyler üstünden yaklaşımı homunkular işlevselcilik olarak tanımlanmıştır. Homunkular işlevselliği zeki bir sistemin, zihnin, sistemin bütününden daha az zeki birçok alt sistemin birlikte işlev görmesiyle açıklanabileceğini savunur. Alt sistemler homunkuli olarak adlandırılır. Homunkuli yaklaşımının, zeki bir ajanın kapasitelerini açıklamak için kullanışlı bir yaklaşım olduğunu ilk öne süren Atteave, homunkuli yaklaşımın işe yaraması için bir alt seviyedeki homunkulusun zeki ajanın kapasitesini birebir taşımaması gerektiğini belirtmiştir. (Lycan, 1991, s: 260) Bir başka ifadeyle zeki bir ajanın performansı, bir alt seviyede işlev gören birden fazla daha az zeki ajanın birlikte işlev görmesiyle açıklanabilir. Bir alt seviyedeki ajanlar, bir üst seviye ile kıyaslandığında daha az zeki ama yine de özerk ajanlardır. Dolayısıyla homunkular işlevselliğe göre bir insanın bilişsel yetileri, bir alt seviyede insana benzer bilişsel kapasiteleri olan ajanların ortak işleviyle açıklanır. Dennett beynin ve beynin alt sistemlerinin de yaşayan bir insanı çağrıştırdığını, hiyerarşik olarak her alt seviyenin biraz daha az zeki özerk ajanlardan oluştuğunu ve nöron seviyesine indiğimizde nöronların bir makine ile yer değiştirilebilecek, açık ya da kapalı durumlarda olabilen özerk ajanlar olduğunu belirtir. Açık ve kapalı olmak üzere iki

durumu olan nöronları ise bir makine ile yer değiştirmek mümkündür. (Dennett, 1978, s: 80-81)

Dennett en temel seviyede makineler ile değiştirilebilecek açık veya kapalı durumda olan nöronların en üst seviyede zihni ortaya çıkardığını savunmaktadır. Dennett, kimya ile fiziği, biyoloji ile kimyayı ve sosyal bilimlerle biyolojiyi birleştirmenin mümkün olduğunu ve öyle olmasının arzu edildiğini belirtir. Dolayısıyla üst seviye bir bilim olan psikolojinin de temelde biyolojiye indirgenebileceğini düşünmektedir. Ancak Dennett, indirmeciliğin üst düzey bilimlerin kavramlarının alt seviye bilimin kavramlarıyla değiştirilmesi olmadığını da vurgular. Düzenlilikler açısından düşünüldüğünde fizik ve kimya kendi içinde düzenliliklere sahip bilimlerdir ve söz konusu olan düzenlilik kendi sınırları içinde değişik kavramlarla ifade edilirler ancak kimya bilimini fizik biliminin temellerinde açıklamak mümkündür. Ancak moleküler düzeyde iş gören kimya bilimini temel parçacıklar seviyesindeki fiziğinin kavram ve yasalarıyla açıklanması mümkünse de kendi düzeyinde bu kavram ve yasalarla iş görmesi düşünülemez. Dennett indirgemeci yaklaşımlar arasında iyi indirgemeciliği ve açgözlü olarak nitelendirdiği indirgemeciliği ayırt eder; açgözlü indirgemecilik sistemlerin karmaşıklıklarını hafife alarak kuramların bir takım aşama ve düzeylerini es geçme eğiliminde olan indirgemeciliktir. Dolayısıyla üst seviye bir bilim olarak tanımlanan psikolojiyi doğrudan nöronlara indirgemeye çalışmak hem beynin yapısının karmaşıklığını hafife almak hem de beyinde değişik düzeylerde gözlemlenen birçok düzenliliği es geçmek açgözlü bir indirgemecilik olacaktır. Dennett'a göre iyi bir indirgemecilik birçok olgunun gizemini yok ederek, onları çok daha emin temeller üstüne yerleştirirken, açgözlü indirgemecilik gerçek olguların varlığını inkar etmemize yol açabilir. (Dennett, 1995, s: 98-100) Dennett bilincin içeriklerinin yani *qualia*'nın fiziksel bir gerçeklik olmadığı düşüncesinin açgözlü bir indirgemecilik yaklaşımın sonucu olduğunu savunur.

Bilimsel yöntem bir fenomendeki epistemik bağlantıları ortaya çıkarmak üstüne kurulmuştur. Bilim sistemdeki epistemik bağlantılar üzerinden ontolojik sınıflandırmalar yapar. Dolayısıyla bilimsel bir yaklaşımda zihni açıklama girişiminin ilk olarak ontolojik sınıflandırmalarla başlaması bilimsel metoda uygun değildir. Dennett, epistemolojik olarak gözlemlediğimiz düzenliliklerin gerçek düzenlilikler olduğunu ve zihnin beynin en üst seviyesinde gözlemlenen düzenliliklerin ortaya çıkmasıyla belirdiğini, bir başka ifadeyle zihnin beynin karmaşıklığı sonucu zayıf belirlenimle ortaya çıkan bir fenomen olduğunu savunmaktadır. Dennett'a göre insan

zihni de en nihai aşamada semantik anlam taşıyan semboller üstünden iş gören bir Turing makinesidir.

Yapay zeka araştırmacıları hedeflerini insan zihninin özelliklerine sahip makineler yapmak olarak belirlemiştir. Dolayısıyla yapay zeka çalışmalarının temelinde zihinsel yetilerin çoklu gerçekleştirilebilir olduğu düşüncesi bulunmaktadır. İnsan zihninin hangi işlevsel bakış açısına göre modellenmesi gerektiği sorusu ise yapay zeka araştırmalarındaki mimarilerin farklılaşmasına neden olmuştur. Ancak tüm yapay zeka çalışmalarında temel amaç belli işlevleri gerçekleştiren makineler üretmektir. Yapay zeka çalışmaları, özellikle üst düzey zihinsel yetileri taklit etmeye çalışırken zihnin üst seviye özellikleriyle uğraşırlar. Bu özellikler ise genelde halk psikolojisinin kavramlarıyla tanımlanan özelliklerdir. Dennett yönelimsel duruşun araçsal bir tutum olduğunu öne sürerek üst düzey zihinsel özelliklerin ontolojik temellerini yapay zeka çalışmaları ve hatta insan zihnini anlama çabaları alanının dışına atmış ve güçlü yapay zeka sistemleri yaratma olasılığını iyi belirlenmiş epistemolojik problemlerin çözümü problemine indirgemiştir. Bir sonraki bölümde yapay zeka araştırmacılarının belirlemiş olduğu problemler ve bu problemlerin çözümüne dair geliştirilen farklı bakış açıları incelenecektir.

# 2. BÖLÜM

# GÜNÜMÜZDE YAPAY ZEKA

## 2.1 YAPAY ZEKA ARAŞTIRMALARININ BAŞLANGICI

Turing'in öne sürdüğü fikirlerin ardından yapay zeka çalışmalarının bilgisayar bilimleri içinde tam teşekkülü bir araştırma sahası olarak kabul görmesi ise art arda yapılan "Öğrenen Makineler Oturumu" (1955 Los Angeles), "Yapay Zeka Yaz Araştırma Projesi" (1956 Dartmounth) ve "Düşünce Süreçlerinin Mekanikleştirilmesi" (1958) toplantılarına tekabül eder. (Nilsson, 2011, s: 75) 1955 yılında yapılan "Ögrenen Makineler Oturum"u toplantısında Clark ve Farley'in sinir hücrelerine benzer öğelerden oluşan ağlardan yapılmış örüntü tanımlama deneyleri hakkında sundukları makale yapay zeka çalışmalarının "yapay sinir ağları" başlıklı alt branşının öncü makalesi kabul edilir.

1956 yılındaki "Yapay Zeka Araştırma Projesi" toplantısında ise Minsky, daha sonra "Yapay Zekaya Doğru" başlığıyla yayımlanacak ve yapay zeka alanının öncü makalesi olarak kabul edilerek birçok çalışmaya temel olacak makalesinin sunuşunu yapmıştır. Bu makalede Minsky yapay zekanın sahip olması gereken becerileri görüntü tanıma, öğrenme, problem çözme, karar verme, plan yapabilme olarak tayin ederek, bu becerilerin matematiksel olarak nasıl ifade edilebileceğini gösteren taslaklar oluşturmuştur. Minsky, ayrıca, eğer bir varlık, hipotetik bir soruyu, deney yapmadan cevaplama yetisine sahipse cevabın ancak varlığın içindeki bir modelden gelebileceğini, konulması gereken hedefin içsel bir modele göre davranan bir makine inşa etmek olduğunu belirtmiştir. (Minsky, 1961, s: 28) Yapay zeka çalışmalarının temelini oluşturan bu düşünce günümüzde dahi geçerliliğini korumaktadır.

Dartmounth toplantısındaki bir başka önemli sunum ise Newell ve Simon'ın, simgesel mantıkla teorem ispatlamaya yarayan "mantık kuramcısı" adlı programı tanıtmalarıdır. Turing'in hesaplama yönteminden temel alan mantık kuramcısı programı "fiziksel simge hipotezi" üstüne kurulmuş ilk yapay zekanın örneğidir. Newell ve Simon sayısal bilgisayarların evrenselliğinin doğru yazılıma sahip olduğu takdirde zekayı mekanikleştirmek için kullanılabileceği düşüncesinden yola çıkmıştır. Çalışmalarının temelini oluşturan "fiziksel simge hipotezi", simge yapılarının dönüştürülmesiyle çalışan sistemlerin zeki eylemler için gerekli ve yeterli imkanlara sahip olduğunu öne

sürer. (Nilsson, 2011, s: 69) Sembolik mantıkla çalışan bilgisayar programları çözülecek probleme göre başlangıç sembolleri tanımlar ve bu sembolleri algoritmanın ön gördüğü şekilde dönüştürerek iş yaparlar. Genel olarak bilgisayar programı bir problemi çözerken sembollerin olası tüm dönüşümlerini hesaplar ancak bazen dönüşüm süreçlerinden bir sonuç ifadesi yaratamaz. Newell ve Simon, mantık kuramcısı programlarında, sadece çözüme açılan yolda olması muhtemel dönüşümleri yaratmak için “sezgisel yönteme” başvurmuşlardır. Bir problemin sezgisel yöntemini, o problemi belki çözecek ama bunun için garanti vermeyen süreç olarak tanımlamışlardır. Sezgisel yöntemler bulmak ve kullanmak yapay zeka çalışmalarında sezgisel programlama alanını doğurmuştur. (Nilsson, 2011, s: 118) Simon, mantık kuramcısının problemleri çözerken insanların kullandığı gibi sezgisel yöntemler kullandığını düşündüğünden, mantık kuramcısını “düşünen makine” olarak adlandırmıştır.

Yapay zekanın bir araştırma alanı olarak tanınmasını sağlayan diğer bir toplantı ise 1958 yılındaki “Düşünce Süreçlerinin Mekanikleştirilmesi”dir. Bu toplantının raporunun ön sözüne göre toplantının amacı, yapay düşünce, karakter ve örüntü tanıma, öğrenme, otomatik dil çeviri, otomatik programlama, endüstriyel planlama ve yazı işlerinin mekanikleştirilmesini araştıran bilimcileri bir araya toplamak olarak belirtilmiştir. (Nilsson, 2011, s: 85) Yapay düşünce, karakter ve örüntü tanıma, öğrenme, otomatik dil çeviri, otomatik programlama, karar verme yapıları günümüzde de yapay zeka araştırmalarının temelini oluşturmaktadır.

İlk başlangıç tarihinden günümüze yapay zeka çalışmaları iki odağa ayrılabilir. İlk odak, Newell ve Simon’ın fiziksel simge hipotezinden yola çıkarak hedefini zeki makineler tasarlamak olarak belirlemiş yapay zeka çalışmalarıdır. Bu çalışmalar problem çözme kabiliyetine sahip bir zekanın ancak semboller üstüne kurulabileceğini ve insan dilinin bir semboller sistemi olduğu düşüncesinden yola çıkarak insan zekasının ancak sembolik bir işleyişe sahip olabileceğini savunurlar. Özellikle yapay zeka alanın öncülerinden McCarthy, “Sağduyulu Programlar” makalesinde yapay zeka araştırmalarının misyonunu zekayı inceleyerek hesaplamanın temel gerekliliklerini açığa çıkarmak olarak tanımlamıştır. Yeni bilgileri çıkarım yaparak öğrenebilen, soyut kavramlar ile işlem yapabilen zeki bir ajanın sahip olması gereken özellikleri inceleyen McCarthy, bu ajanın insan dili gibi sembolik bir yapıya sahip olması gerektiği ve makinedeki bilgi temsillerinin sembolik ifadeler olması gerektiği sonucuna varmıştır**.** (McCarthy, 1959, s: 3)

Yapay zeka çalışmalarının diğer odağı ise yapay sinir ağları olmuştur. İnsan beyninin nöron hücrelerinden oluştuğu ve insan beyninin bu hücrelerin birbiriyle etkileşimleri ile hesap yaptığı tezinden yola çıkan yapay sinir ağları araştırmaları hedefini insan beyninin yapısını taklit ederek insan gibi düşünen makineler yaratmak olarak belirlemiştir.

Bu iki odak yapay zeka çalışmaları tarihi boyunca birbirileriyle çatışmıştır. Felsefe dünyası da yapay sinir ağlarını eleştirmiş ve insan zihninin semboller ile iş gördüğünü öne sürmüştür. Tarihsel olarak değişik dönemlerde yapay sinir ağları yapay zeka çalışmalarının ağırlık noktasını oluşturmuş, bazı dönemlerde ise sembolik mantığa dayanan çalışmalar ön plana çıkmıştır. Günümüzde hala sürmekte olan bu tartışmada ibre yapay sinir ağlarından yana ağırlık kazanmış olsa da yapay sinir ağlarının karşılaştığı en ciddi sorun soyut düşüncenin yapısındaki sembolik mantığın mekanikleştirilmesidir. Günümüzde yapay zeka çalışmalarının sembolik mantığı yapay sinir ağlarına adapte edebilecek hibrit sistemler arayışına yoğunlaştığı söylenebilir.

## 2.2 PERCEPTRON

Yapay sinir ağlarının gelişimine bakıldığında, tarihin ilk öğrenebilen makinesi olan "Perceptron"un 1959 yılında Rosenbaltt tarafından inşa edildiği görülür. Psikoloji alanında uzmanlaşmış olan Rosenbaltt, yapay sinir ağlarına insanın biliş, öğrenme ve bellek yetilerinin olası bir modeli olarak ilgi duymuştur. (Rosenblatt, 1958, s: 386) Perceptronun alt yapısını oluşturan iki düşünceden ilki 1943 yılında yayınlanan McCulloch-Pitts'in, "A Logical Calculus of the Ideas Immanent in Nerveous Activity" başlıklı makalesinde öne sürdüğü bir fikirdir; bu fikir nöronların birbirleriyle etkileşiminin temel mantık devreleriyle açıklanabileceğine dairdir. İkincisi ise 1949 yılında Donald Hebb tarafından ileri sürülen öğrenme kuralıdır; bu kural nöron ağlarının nasıl öğrendiğine ve nöron ağlarının bilgiyi yapılarında nasıl muhafaza ettiğine dairdir. Bu kural literatüre "Hebb öğrenme kuralı" olarak geçmiştir.

Temel sinir hücreleri olan nöronları diğer hücrelerden ayırt eden başlıca özellikleri bilgi taşımalarıdır. Nöronlar merkezi sinir sisteminin temel yapı taşlarını oluşturmakla beraber tüm vücudumuzda bulunmakta ve vücudun içsel ve dışsal duyumlarını beyne iletmenin yanı sıra lokal olarak vücudumuzun işlemesinde de aktif rol oynamaktadırlar. Nöron hücreleri yapısal ve işlevsel olarak farklılıklar gösterseler de bir nöron genel

olarak “soma” olarak adlandırılan ve hücrenin kendi yaşamsal fonksiyonlarını yürüten hücre gövdesinden, bilgi taşıma işlevini yerine getiren “akson” uzantısından ve nöronun diğer nöronlarla iletişiminde bilginin algılanmasını sağlayan “dendirit” uzantılarından oluşur. Nöronlar bilgiyi elektriksel sinyaller ile taşırken nöronlar arası iletişim kimyasal bir mekanizmaya sahiptir. (Crick, 1997, s: 103)

Bir nöron uyarıldığında uyarının akson boyunca iletimi, hücre içi ve hücre dışı sıvıların içinde bulunan iyonların akson boyunca yer alan ve kontrollü şekilde açılıp kapanabilen kanallar aracılığıyla yer değiştirmesiyle gerçekleştirilir. Nöron hücresinin içindeki sıvıda bolca iyonlaşmış protein molekülleri ve potasyum iyonları (K+), hücre dışı sıvıda ise ağırlıklı olarak klorür (Cl-) ve sodyum iyonları (Na+) bulunur. Dinlenme halinde uyarılmamış bir nöron, iç ve dış sıvıların oluşturduğu yük farkı sonucu dış ortamına göre ortalama -70 Voltluk bir potansiyel farka sahiptir. Nöron uyarıldığında, aksonun o bölgesindeki sodyum kapıları açılır ve hücre dışı sıvıdaki sodyum iyonları hücre içine aktarılarak, hücre içindeki potansiyel farkı azaltılır. Bu durum aksonun o bölgesindeki potasyum kapılarının açılmasına ve hücre içindeki potasyumun dışarı pompalanmasına neden olur. Bu süreç ise aksonun o bölgesindeki potansiyel farkı yeniden dinlenme potansiyeline döndürür. Sodyum ve potasyum kapılarının artarda açılıp kapanmasıyla oluşan elektrik sinyali akson boyunca taşınır. (Crick, 1997, s: 107)

İki nöron arasındaki iletişimin gerçekleştiği bölgeye sinaps denmektedir. Sinapslar, elektriksel ve kimyasal sinapslar olarak ikiye ayrılırlar. İki tip sinapsta da iletişim kimyasal bir yapıdadır. Elektriksel sinaps oluşturmuş nöron hücrelerinin hücre zarları birbirlerine çok yakındır ve iletim hücreler arası seçici olmayan protein kanalları olan *connexon*ların açılıp, kapanması aracılığıyla gerçekleşir. Kanallar açıldığında iyonlar bir nörondan diğerine geçerek sinyalin iletilmesini sağlar. Elektriksel sinapslar çift yönlü çalıştıkları gibi, bilgi iletim hızları kimyasal sinapslara göre oldukça hızlıdır. Beynimizin bilişsel yeteneklerinin genel olarak kimyasal sinapslarla iletişim kuran nöronlar tarafından yürütüldüğü düşünülmektedir. Kimyasal sinapslarda nöron hücreleri arasındaki iletişim nöro-iletici denen kimyasallar ile gerçekleşir. İnsan beyninde yirmiden fazla nöro-iletici bulunduğu bilinmekle beraber tüm nöro-ileticilerin nöronların iletişimi üstündeki etki mekanizmaları tam olarak açıklığa kavuşturulamamıştır.

Kimyasal sinaps oluşturan iki nörondan sinyali gönderecek pre-nöron aksonunun ucunda içleri nöro-ileticiler bulunan torbacıklar, sinyali alacak olan post-nöronun dendritinde ise nöro-ileticileri algılayan dikenler bulunmaktadır. Aksonda iletilmekte olan sinyal akson uçuna ulaştığında değişen voltajla torbacıklar açılır ve nöro-ileticiler

serbest kalır. Serbest kalan nöro-ileticiler, dendrit dikenlerinde bulunan reseptörler tarafından algılanır. Pre-nöronun, post-nöronu uyarması, sinapstaki nöro-ileticilerin türleri ve yoğunluklarıyla belirlenen sinapstik güce bağlıdır. Bir nöronun uyarılabilmesi için gereken uyarım miktarına uyarılma eşiği denir. Eğer uyarılma eşiği aşılırsa post-nöron ateşlenir, eğer uyarı eşiği aşılamazsa post nöronda herhangi bir aktivite gözlemlenmez. Bu noktada nöronların her seferinde tek bir sinyal taşımaktan daha çok art arda sinyaller taşımaları önemlidir. Taşınan sinyalin frekansı, kısaca nöronun belli bir zamanda art arda ateşlenme sayısı, sonraki nöron uyarılmasında önemli bir faktördür. Pre-nöronun taşıdığı sinyalin frekansı ile sinapsın gücünün çarpımından, pre-nöronun post nöronu uyarma şiddeti bulunur. Bu işlemin temelinde akson ucunun salgıladığı nöro-iletici miktarının bulunması yatar. Beyinde uyarıcı ve baskılayıcı olmak üzere iki çeşit sinaps bulunur. Uyarıcı sinapslar sinyalin bir nörondan diğer nörona iletilme olasılığını arttırırken, baskılayıcı sinapslar sinyalin bir nörondan diğer nörona iletilme olasılığını azaltırlar. (Crick, 1997, s: 110-116)

Mc-Culloch ve Pitts, nöronların temel mantık devrelerini oluşturabildiğini göstermiş ve nöron kalkülüsünü oluşturmuşturlar. Temel mantık işlemleri; “Ve”, “Veya” ve “Değil” kapılarından oluşur. İki pre-nöron, A ve B’nin, bir post nöron C’ye bağlandığı bir yapı düşünüldüğünde, eğer A ve B nöronlarının uyarma güçleri tek başına C nöronunu ateşlemeye yetmiyorsa ama A ve B nöronlarının uyarma güçlerinin toplamı C nöronunu ateşleyebiliyorsa, bu sistem bir “Ve” kapısıdır. Aynı sistemde, eğer tek başına A veya B nöronunun uyarma güçleri C nöronunu ateşleyebiliyorsa bu sistem bir “Veya” kapısı oluşturur.  Değil kapısı ise baskılayıcı sinaps olarak düşünülür ve nöron aritmetiğinde eksi işaretiyle belirtilir. Gerçek bir nöronun diğer nöronlarla birden fazla sinapstik bağlantı oluşturduğu bilinmektedir, her bağlantı noktasında, bir mantık kapısı oluşturmak mümkündür. (McCullock, Pitts, 1948, s: 102-113) Hesaplamalı zihin kuramında mantık kapıları kullanılarak beynin birçok faaliyeti açıklanabilir ve bu tarz devreler, klasik bilgisayar devrelerini taklit edebilirler. Ancak nöron devrelerinin, klasik bilgisayar mimarisinde bulunmayan önemli özelliği bulanık mantık devreleri oluşturabilmeleridir.

Yapay sinir ağlarının temelini oluşturan bir başka nörobilim temelli düşünce ise Hebb öğrenme kanunu olmuştur. 1949 yılında Donald Hebb, “Davranışın Organizasyonu” adlı kitabında, düşünmek, öğrenmek gibi soyut zihinsel kavramların, beyindeki mekanizmalar ve “hücre toplulukları” ile açıklanabileceğini savunmuştur. Hebb, genel olarak iki ya da daha fazla hücreden oluşan bir sistemde, aynı anda aktive olan

hücrelerin aralarında bir ilişki oluşturmasının ve aynı sistem içindeki hücrelerden birinin aktive olmasıyla diğer hücrelerinde aktive olmasının kabul edilen düşünceler olduğunu belirmiştir. (Hebb, 1949, s: 52) Aynı şekilde, birbiri ardına ve hatırı sayılır bir zaman süresince aktive edilen nöronların hücre toplulukları meydana getirerek hafızayı ve ardışık düşünme süreçlerini oluşturabileceğini, dolayısıyla bir uyarana tepki veren mekanizmaların birbirleriyle ilişkili nöron topluluklarıyla açıklanabileceğini öne sürmüştür. Hebb'e göre öğrenme, birbiri ardına ateşlenen nöronların belli fiziksel ve kimyasal değişimler sonucu birbiriyle kalıcı olarak ilişkilenmesinin bir sonucudur. Günümüzde Hebb öğrenme kuralı olarak bilinen bu süreç, "birlikte ateşlenen nöronlar, birbirine bağlanır" şeklinde özetlenebilir. Hebb, birlikte ateşlenen nöron toplulukları arasında güçlü ve kalıcı sinaptik bağlantılar oluştuğunu, kullanılmayan bağlantıların ise zaman içinde kaybolduğunu belirtmiştir. Hebb öğrenme kanunu, öğrenme sürecine nöronal seviyede bir açıklama getirirken hafızanın oluşumunu da açıklamıştır. Sinapstik bağlantıların güçleriyle çizilen nöronal aktivasyon deseni, tekrarlayan işlemlerle, yerine getirilen işlev için dinamik bir harita oluşturmaktadır.

Hebb, soyut zihinsel işlevleri, nöronlar arasında kurulan ilişkilerle açıklayarak, günümüzde "bağlantıcılık" olarak bilinen ve yapay sinir ağları çalışmalarının üstüne kurulduğu görüşün doğmasını sağlamıştır. Bağlantıcılık, insanın bilişsel yetilerinin tamamen nöron ağları ve bu nöron ağları arasında kurulan bağlantıların özellikleriyle açıklanabileceğini savunur. Dolayısıyla bağlantıcılık zihinsel yetileri gerçekleştirenin doğrudan sistemin yapısındaki nedensel ilişkiler olduğunu düşünmektedir.

Hebb öğrenme kanunun bir uygulaması niteliğinde olan Rosenblatt'ın *perceptronu* girdi ve çıktı düzlemi olmak üzere iki düzlemlerinden oluşmaktaydı. Bu düzlemlerdeki her düğüm noktası bir nörona tekabül ediyordu ve her nöron McCulloch ve Pitts'in nöron işlemlerini yapacak devreler oluşturarak birbirleriyle bağlantılıydı. Örnek olarak, eğer *perceptron*un el yazısıyla yazılmış sıfırdan dokuza kadar olan rakamları okuması isteniyorsa, girdi katmanında sayının yazıldığı ekranın piksel sayısı kadar girdi düğümü bulunuyordu. Çıktı için ise olası cevap sayısıyla aynı sayıda nöron içeriyordu, dolayısıyla rakam okuyan *perceptron*un on çıktı düğümü bulunuyordu. *Perceptron* girdilere göre doğru çıktıyı verecek şekilde eğitiliyor, bu eğitim ise Hebb'in öğrenme kanununu esas alıyordu. *Perceptron*da her düğüm bir aktivasyon fonksiyonuyla tanımlanıyordu. Her girdiye tekabül eden doğru sonucu alınırken ise düğümler arasındaki ağırlık katsayısı ayarlanıyordu. Bu şekilde makine sadece girdiler ve sonucun doğruluğundan yola çıkılarak düğümleri arasındaki ağırlıkları kendiliğinden

ayarlayabiliyor, dolayısıyla her girdi için doğru çıktıyı vermeyi öğreniyordu. Makinede düğümler arasındaki ağırlıklar makinenin hafızasını oluşturmaktaydı. Ağırlıklar son hallerine geldiğinde *perceptron* el yazısı ile yazılan rakamları yüzde doksan sekizlik bir başarı ile tanıyabiliyordu, dolayısıyla *perceptron* el yazısı okumayı öğrenmeyi başarmıştı.

Rosenblatt'ın *perceptron*u dışında, atmışlı yılların sonuna dek birçok yapay sinir ağı projesi yapay zeka çalışmalarında ön plana çıkmıştır. Ancak Minsky 1969 yılında *perceptron* ve genel olarak yapay sinir ağları mimarilerinin sorunlarına dikkat çeken bir makale yayınlamıştır. Bu makalede *perceptron*un en ciddi kusurunun özel veya kapısına sahip olmamasını göstermiştir. A özel veya B kapısı, A veya B ama ikisi birden değil anlamına gelir. Eğer A açıksa ve B açıksa sonucun kapalı olması gerekir, *perceptron* düğümler arasındaki ağırlıklarına göre karar verdiğinden bu işlemi yapmayı başaramaz. Eğer A'nın ve B'nin açık olduğu durumlar için düğümler arasındaki bağlantı güçleri azaltılırsa A'nın ve B'nin kapalı olduğu durumda açık olması gerekirken sonuç kapalı olur. (Pinker, 2016, s: 135-136) Özel veya kapısını doğru şekilde temsil etmenin tek yolu ise iki katman arasına saklı bir katman eklemektir. Ancak bu sefer karşılaşılan sorun ara katmandaki düğüm ağırlıklarının ayarlanmasıdır. Minsky, çok katmanlı *perceptron*larda Hebb öğrenme kuralının uygulanamayacağını iddia etmiştir. (Minksy, 1969, s: 240)

Minsky'nin yapay sinir ağlarını eleştiren makalesinin ve dilin sembolik bir yapısı olduğu görüşünün etkisiyle 1980'li yılların sonlarına dek yapay zeka çalışmalarının odağı yapay sinir ağları çalışmalarından fiziksel sembol hipotezi üstüne kurulmuş, sembolik mantıkla çalışan sistem araştırmalarına kaymıştır.

## 2.3 SEMBOLİK SİSTEMLER

Sembolik mantıkla çalışan sistemlerin temelinde insan dilinin sembolik bir yapısı olduğu ve insan zihninin sembolleri dönüştürerek iş gördüğü görüşü yatar. Sembolik sistemler aynen evrensel Turing makinesindeki gibi sembol dönüşümlerini sıralı, bir başka deyişle seri bir algoritma üstünden gerçekleştirir.

Fiziksel sembol hipotezi üstüne kurulmuş araştırmaları yönlendiren en önemli eser Newell ve Simon tarafından 1972 yılında yayınlanan *Human problem solving* kitabıdır. Newell ve Simon bu kitapta öncelikle insanın bilişsel yeteneklerini bilinçli ve bilinç dışı

olarak ikiye ayırmış, problem çözme yetisinin tamamen bilinçli bir süreç olduğunu belirtmiştirler. Daha sonra, problem çözme görevlerini yerine getiren insan deneklerle deneysel çalışmalar gerçekleştirmişlerdir. Bu çalışmalar, göz hareketlerinin izlenmesinden, soyut problem çözen bir uzmanın, problem çözme aşamasında sesli olarak ifade ettiği düşüncelerinin incelenmesine dek geniş bir yelpazeyi kapsamıştır. Newell ve Simon, yapılan deneyleri inceleyerek, insanların problem çözme yetilerinin belli bir bilişsel mimari ile modellenebileceğini sonucuna ulaşmışlardır. Newell ve Simon'ın modellemesinin temeli, olgulardan oluşan açık temsillere belli kuralların seri şekilde uygulanmasına dayanmaktadır. (Newell, 1972, s: 53-71)

Newell ve Simon'ın insanın problem çözme yetisi bütünlüklü bir araştırmaydı, değişik düşünme süreçlerinin yanı sıra hafıza üstüne de araştırmalar yapılmıştı. Modellemede ise problemin kısımlarını tanımlayan düzlemler grafiksel yapılara sahipti, düzlemlerin çakıştığı düğümler bilgi durumlarının sembolik gösterimini sunuyordu. Çalışan algoritma düğüm noktalarındaki sembollere göre yaklaşık her beş saniyede bir, operatör öbekleriyle sembolleri manipüle ediyordu. Bu operatör öbeklerine "üretim sistemleri" denmekteydi. Aynı zamanda Newell ve Simon'ın modelinde üç ayrı hafıza sistemi bulunmaktaydı. Dış hafıza sistemi problemi tanımlayan, problemin sınırlarını belirten çerçevedeki bilgiler için tasarlanmıştı. İnsan zihninden esinlenmiş diğer iki hafıza birimi ise kısa süreli hafıza birimi ile uzun süreli hafıza birimiydi. Uzun süreli hafıza yapılacak görevlerin kurallarını tutuyordu. Kısa süreli hafıza ise yerine getirilecek görev hakkındaki dinamik bilgiden oluşuyordu. Kurallar, genel olarak "Eğer" kodlarıyla yazılmıştı, kısa süreli hafızadaki dinamik bilgilerin sağladığı koşulla göre program harekete geçiyordu. Bu hareket, öncelikle kısa süreli hafızadaki bilginin değiştirilmesiyle başlıyor ve göreve göre belli algoritmaların çalışmasını içeriyordu. Çalışan algoritmaların sonucu ise uzun süreli hafızaya eklenerek, uzun sürekli hafızadaki bilginin yenilenmesini sağlıyordu. Ancak uzun süreli hafızadaki temel kurallar listesi değişime uğramıyordu. (Young, 1999, s: 224-231)

1980'li yıllara dek yapılan yapay zeka araştırmalarında ağırlıklı olarak Newell ve Simon'ın yöntem ve mimarileriyle oluşturulmuş uzman sistemler ön plana çıkmıştır. Ancak sembolik mantık kullanılarak yapılan çalışmalar elde ettikleri başarılara rağmen yapay zekaya dair beklentileri karşılamaktan uzak kalmışlardır. Newell Simon mimarileri günümüzde "eski moda iyi yapay zeka" olarak adlandırılmaktadır.

## 2.4 YAPAY SİNİR AĞLARI

Yapay zeka çalışmalarının ağırlık odağının sembolik sistemlerden yapay sinir ağlarına yönelmesinde beynin çalışma prensiplerine dair modeller oluşturulmasının önemli etkisi olmuştur. Yapay zeka çalışmalarının başlangıcından günümüze beynin çalışma prensipleri hakkındaki görüşler köklü değişikliklere uğramıştır.

### 2.4.1 Beynin Yapısı

İnsan beyni evrende bilinen en karmaşık sistemdir ve insan beyni yaklaşık yüz milyar kadar nörondan oluşur ve her nöron diğer nöronlarla yaklaşık on ila on bin bağlantı yapar. Makro ölçekte dıştan beyne bakıldığında beynin birbiriyle hemen hemen simetrik iki yarım küreden oluştuğu gözlemlenir. Selebral korteks olarak adlandırılan bu iki yarım küre birbirlerine korpus kallosum adı verilen kalın bir sinir lifi ile bağlıdır. Bu sinir lifi iki yarı kürenin iletişimini sağlayarak, iki yarım kürenin birbirinden haberdar olarak eşgüdümlü işlev görmesini sağlar. Beynin dış yüzeyini oluşturan korteks (beyin kabuğu), yaklaşık 3 mm kalınlığında, kıvrımlı bir yapıdır ancak koku dışındaki duyumların algılanması, düşünme, ilişkilendirme, karar verme, dikkat, öğrenme, hafıza ve bilinçli hareketler gibi birçok üst düzey fonksiyondan sorumludur. Beyinde bulunan yüz milyar nöronun otuz milyarı kortekste bulunur ve korteks nöronları, özellikle primadial nöronlar, diğer bölgedeki nöronlarla kıyaslanmayacak kadar çok sinaptik bağlantıya sahiptir. Korteksteki nöron hücreleri incelendiğinde elliden fazla farklı şekilde nöron olduğu tespit edilmiştir. (Edelman, Tononi, 2019, s: 62)

Beyin kabuğu evrimin son aşamasında gelişmiş olduğundan bu kabuğa, neokortekste denmektedir. Korteksin kıvrımlı yapısının amacı yüzey alanını genişletmektir. Bu kıvrımlara girus, daha derin yarık görünümündeki kıvrımlara ise sulkus denmektedir. Sulkuslar, anatomik olarak beyni frontal, pariyetal, oksipital ve temporal lob olarak dört ana loba böler. Frontal lobun ön bölümünde yer alan prefrontal korteks insan beynini diğer memelilerin beyninden ayıran en önemli bölgedir. Prefrontal kortekste oluşan hasarlar kişilik değişimine neden olur. Prefrontal korteks, ahlak ve toplumsal kuralların algılanması, irade ve uzun vadeli hedeflere göre hareket etmek gibi işlevlerden sorumludur. (Damasio, 2018, s: 62) Beyinde temporal lobun altında kalan bölüm ise yakın zamanda insular lob (İnsula) olarak tanımlanmış, bu lobun vücudun içsel durumlarını ayırt ederek susuzluk, terleme, huzursuzluk gibi duyguları ortaya çıkardığını anlaşılmıştır. (Damasio, 2018, s: 111)

Beyin kabuğunun altında evrimsel olarak daha eski yapılar bulunur. Bazal ganglia korteks altı bir yapıdır ve talamusunda dahil olduğu birçok paralel işleyen çekirdek yapısından oluşur. İstemli hareketlerden, göz kırpma hareketine dek tüm motor hareketlerin modülasyonu bazal ganglianın görevidir. Ayrıca korteks altında koku duyumundan sorumlu olan oldfactör yapısı vardır. Beyindeki diğer korteks altı yapılardan -amigdala, hipokampüs, hipotalamus ve bazal ganglia sisteminde de görev alan talamus- limbik sistemi oluştururlar. Amigdala tehdit içeren uyaranların tanımlanmasında rol oynayarak “savaş ya da kaç” tepkisini oluşturur. Aynı şekilde çevresel pozitif uyaranları da işleyen amigdalanın bilinçli hafıza oluşumunda önemli bir yeri vardır. Simetrik olarak iki yarım kürede de bulunan talamus beyin içinde beyin olarak adlandırılan bir yapıdır. Koku duyusu hariç tüm duyusal bilgi ilk olarak talamusta değerlendirilerek korteksin uygun bölgelerine gönderilir. Hipokampüs, hipokampal formasyon denilen bir yapının parçasıdır. Hipokampüs korteksten aldığı bilgileri işleyerek kortekse, talamusa ve hipotalamusa bilgi gönderir. Hipokampüs yön bulmada ve bilinçli anıların oluşumunda kilit bir rol oynar. Hipotalamus ise beyin sapının hemen üstünde bulunur. Çeşitli çekirdeklerden oluşan hipotalamus vücut ısısı, kan basıncı, nabız hızı gibi hayati fonksiyonların kontrolünü üstlendiği gibi vücuttaki hormonların salınımı ve salınım miktarlarını ayarlar. Beyin sapı omurilik ile beynin iletişimi sağlayan yapıya denmektedir.

Kortekste bazı bölgelerin belli işlevlerde uzmanlaştığı fikri Broadman’ın çalışmalarına dayanmaktadır. Broadman korteks yapısını ayrıntılı bir şekilde inceleyerek korteksi elli iki bölüme ayırmış ve her bölümün ayrı bir işlevi olduğunu iddia etmiştir. Beyinde belli duyusal alanların belli işlevlerde uzmanlaştığı bilinmektedir. Beynin belli işlevleri sadece o işlev için özelleşmiş bölgelerde yerine getirdiği düşüncesi modüllerleşme olarak tanımlanır. Genel olarak bu görüş çeşitli beyin hasarlarının aksattığı bilişsel işlevlerin istatistiki olarak incelenmesiyle temellenmiştir. (Caramazza, 1986, s: 41-66)

Vernon Mountcastle beyin korteksinin homojen bir yapısı olduğunu, modülerleşmenin o bölgeye aktarılan girdilerle oluştuğunu öne sürmüştür. Beynin bir bölgesi zarar gördüğünde beynin başka alanlarının o bölgenin işlevini devralması bu görüşü desteklemiştir. Belirli bölgelerin, belirli işlevlerde uzmanlaşmasında önemli bir faktör beynin duyusal bilgileri işlerken topografik haritalandırma yapmasıdır. (Mountcastle, 1978, s: 15-37) Sinir sisteminin organizasyonunda duyusal reseptörlerin komşuluk ilişkileri ile kortekste duyusal girdiyi işleyen hücrelerin komşuluk ilişkilerinin birebir örtüşecek şekilde kurulması topografik haritalandırma olarak tanımlanır. Örneğin

görsel sistemdeki retionotopik haritalandırmada retinada komşu olan reseptör hücreler ile görsel bilginin kortekste işlendiği bölümdeki hücreler aynı şekilde komşuluk ilişkilerini koruyacak şekilde birbiriyle bağlantı kurarlar. İki reseptör hücre retinada birbiriyle komşu ise kortekste bu iki reseptör hücreden bilgi alan nöron hücreleri aynı komşuluk ilişkisine sahiptir. Söz konusu komşuluk ilişkileri bilgi alt düzey hesaplamalar yapan bölgelerden, en üst düzeyde hesaplama yapan bölgelere aktarılırken de korunur. İnsanın sadece görsel sisteminde elliden fazla topografik harita belirlenmiştir. Tüm duyusal reseptörler ve beyinde bu duyusal bilginin işlendiği bölgelerde aynı haritalanma organizasyonu bulunmaktadır. (Churchland, 2019, s: 140)

Mountcastle'ın araştırmalarında belirttiği bir başka önemli nokta ise beynin hesaplama birimlerinin kolonlar olduğudur. Korteksin tek biçimli bir nöronal dokudan oluşmadığı, dikey düzlemde altı tabaka olarak sınıflandırılabildiği boyama teknikleriyle gözlemlenir kılınmıştır. Özellikle duyusal bilgiyi işleyen bölgelerde bu tabakalar ve fonksiyonları incelendiğinde en dış tabakanın diğer tabakalardan gelen nöronların dendritlerinden oluştuğu gözlemlenir. İkinci ve üçüncü tabaka benzer şekilde birbirileriyle bağlanmış nöronları barındırır. Dördüncü tabaka nöronları talamustan sinyal alır, duyusal sinyaller kortekse dördüncü tabakadan girer, bu tabaka ilk ve en basit hesaplamaların yapılmaya başlandığı tabakadır. Dördüncü tabaka işlediği sinyalleri ikinci ve üçüncü tabakaya projeksiyon eder. İkinci ve üçüncü tabaka bilgiyi işlendikten sonra beşinci tabakaya ve yatay düzlemde korteksin diğer bölgelerine projeksiyon yapar. Beşinci tabaka bazal ganglia ve korteksin diğer bölgeleriyle iletişim halindedir. Son olarak beşinci tabakadan sinyal alan altıncı tabaka işlenmiş bilgiyi, tekrar talamusa göndererek döngüyü tamamlar. (Mountcastle, 1997, s: 702)

Korteksteki altı tabakalı bu yapı dikey kolonlar halinde gruplaşmıştır. Her kolon belli hesaplamaları yapan nöron devrelerinden oluşur. Bir kolonda ortalama yüz ila bin arasında birbirileriyle sıkı bağlantılar kurmuş nöronlar bulunur. Kolon içindeki nöronlardan biri ateşlendiğinde kolondaki diğer nöronlar da ateşlenir, her kolondaki nöronlar tek bir işlevi yerine getirmek üzere birlikte aktive olurlar. Aktive olmuş bir kolonun çevresindeki kolonları aktive etmesi uzak bölgedeki kolonları aktive etmesinden daha olasıdır. Genel olarak bir kolondaki nöronlar, komşu kolondaki nöronlarla da bağlantılıdır. Ancak beyindeki yol olarak adlandırılan sinyal aktarım yolları ile belli bir kolon aktivasyonu beynin uzak bölgelerinde de aktivasyona yol açabilir. (Edelman, Tononi, 2019, s: 68) Belli alanların sadece belli duyusal bilgileri

işlemesine işlevsel ayrışma denmektedir. (Edelman, Tononi, 2019, s: 66) Bir duyusal bilginin işlenmesi beynin ayrı bölgelerinde gerçekleşir, daha sonra ayrı işlenen bilgi bütünleştirilir. Bütünleşme sürecinde geniş bir ölçekte çok sayıda nöronal grup içeren kortikal alanlar birbirleriyle karşılıklı bağlantı kurarlar. Bu karşılıklı bağlantılar uzantı veya projeksiyon olarak adlandırılır. (Edelman, Tononi, 2019, s: 68) Birleştirme ilk olarak aynı çeşit duyusal bilgiyi, daha sonra değişik duyulara ait bilgileri, örneğin aynı anda işlenmekte olan görsel ve işitsel bilgileri ve bu bilgilerin algılanması esnasında yerine getirilmesi gereken motor hareketleri yönlendiren haritaları dinamik şekilde birleştirir. Her yeni algı ve harekette bu haritalar en uygun sonucu verecek şekilde yenilenmekte ve birleştirilmektedir. Haritalar arasındaki yakın ve uzak bağlantıların ileri geri sinyalleşmesine "yeniden giriş" denmektedir. (Edelman, Tononi, 2019, s: 69) Yeniden giriş, beynin karşılıklı biçimde bağlantılı alanlarının tekrarlayan kesintisiz bilgi alışverişidir.

### 2.4.2 Evrişimli Sinir Ağları

Beynin duyusal bilgileri nasıl işlediği, nöron devrelerinin nasıl iş gördüğü, beynin bilişsel işlevleri nasıl yerine getirdiği günümüzde dahi tam olarak anlaşılamamış olsa da Hubel ve Wiesel'in gerçekleştirdiği ve daha sonra kendilerine Nobel ödülü kazandıran kedi deneyi, beynin çalışma prensiplerinin anlaşılmasında bir dönüm noktası olmuştur. Hubel ve Wiesel gerçekleştirdikleri deneyde bir kedinin birincil görsel korteksinde ateşlenen her nöronun tek tek seslerini kaydetmeyi başarmışlardır. Deneyde kedinin beyninin birincil görsel alanına elektrotlar yerleştirilmiş ve kediye ekrandaki hareketli bir nokta izletilmiştir. Deney esnasında bazı nöronların gösterilen nokta retinada düz bir çizgide ilerlerken, bazı nöronların sadece hareket eden noktanın yönü değiştiğinde ve bazı nöronların nokta sadece belli bir açıda hareket ederken ateşlendiğini gözlemlenmişlerdir. Deney sonucunda Hubel ve Wiesel görsel alandaki nöronlardan her birinin belirli özelliklere göre ateşlendiği, dolayısıyla her nöronun bir özellik seçici olarak işlev gördüğü sonucuna varmışlardır. (Hubel, Wiesel, 1962, s: 106-123)

1980 yılında Fukushima tarafından geliştirilen *Neocognition* doğrudan Hubel ve Wiesel'ın özellik seçici nöron fikrinden ilham alarak tasarlanmış bir yapay sinir ağıdır. Fukushima görsel sistemin en basit özelliklerden başlayarak hiyerarşik bilgi işlemesini doğrudan yapay sinir ağlarında uygulanabileceğini göstermiştir.

*Perceptron*a el yazısı ile yazılmış rakamları sınıflandırılması öğretilirken girdi katmanı görselin piksellerinden doğrudan veri almaktaydı. Fukushima, piksellerdeki bilginin belli özelliklerinin belirlenmesinin *perceptron*un başarısını arttıracağını ön görmüştür. Fukushima, *perceptron*a direkt piksellerdeki veriyi vermektense, ilk olarak girdi görselini özellik seçici nöronlar gibi davranan küçük boyutlu filtreler ile taratmıştır. Söz konusu filtreler görsel sistemdeki alıcı alanlar olarak görev yapmakta ve görseli tararken sadece belirlenmiş belli özellikleri bulmaktaydılar. Örneğin bir filtre görseldeki kenar çizgilerini, bir başka filtre ise görseldeki çizgilerin açılarını belirlemekteydi. Bu tip bir yapay sinir ağında karşılaşılan sorun ise alıcı alanları görselin piksellerinden çok daha küçük boyutta tasarlanan filtrelerin bulduğu özelliklerin esas görseldeki konumun korunması olmuştur. Fukushima'nın bilgilerin konumun korunması için beyindeki topografik harita düzenini andıran bir yapı ile filtrenin bulduğu özellikleri aynı boyutta bir aktivasyon katmanına, filtrelerin tarama yaptığı düzeni koruyarak aktarmıştır. (Fukushima, 1980, s: 193-202) Bu aktarım, kullanılan filtrelerin boyutuna bağlı olarak belli seviyede bilgi kaybına neden olmuş olsa da *perceptron*un rakam tanıma başarısını hatırı sayılır bir seviyede arttırmıştır. Fukushima'nın kullandığı tasarım matematikteki evrişim fonksiyonunu andırdığından Fukushima'nın tasarımındaki gibi hiyerarşik olarak görseldeki özellikleri bularak iş gören yapay sinir ağlarına evrişimli sinir ağları denmiştir.

Geri yayılımın popülerleşmesi ile Fukushima'nın temelini attığı evrişimli sinir ağları geliştirilmiş, 1989 yılında LeCun tarafından el yazısı ile yazılmış posta kodlarını başarıyla sınıflandıran evrişimli sinir ağları gündelik hayatta kullanılmaya başlanmıştır. (LeCun, 1989, 541-551) Bu ağda, filtreler ile görseldeki kenarların, sınırların ve diğer basit özelliklerin belirlenmesi, bu özelliklerin birleştirerek daha üst özelliklerin çıkarılmasından sonra bilgi, gereken sınıflandırılmanın yapılması için çok katmanlı bir yapay sinir ağına aktarılıyordu. LeCun'ın tasarladığı evrişimli sinir ağı rakamları başarıyla sınıflandırmasını sağlamıştır.

Üst düzey bilgiyi işleyen nöronların seçici nöronlar gibi davranıyor olduğu savı, Hubel ve Wiesel'in kedi deneyinin sonucu oluşturulmuş bir savdır. Bir insan yüzünü tanırken o insanın yüzüne ait olan bir nöron olduğunu ve yüz tanımanın bu nöronun ateşlenmesiyle gerçekleştirdiğini savunan bu sav, "babaanne nöronu savı" olarak bilinir. (Churchland, 2019, s: 131) Yapılan bazı açık beyin ameliyatlarından beynimizde tek bir insana ayrılmış nöronlar olduğu keşfedilmiştir. Örneğin açık beyin ameliyatı esnasında bir erkek hastanın beyninde Halle Berry'nin fotoğrafıyla, hatta

değişik Halle Berry temsilleriyle aktive olan tek bir nöron olduğunu gözlemlenmiştir. (Koch, 2012, s: 65) Ancak beyindeki basit işlemlerde seçici nöronların olduğunun anlaşılması, beynin karmaşık işlevleri gerçekleştirirken de aynı şekilde seçici nöronları kullandığı anlamına gelmez. Sadece seçici nöronlar üstüne kurulmuş bir model, beynin birçok işlevini açıklamakta yetersiz kalır. Bilginin fonksiyonel birleşenlerinin beynin bir bölgesinde lokalize olmasına rağmen genel olarak bilginin, tek bir nöronun aktivasyonundan daha çok beyinde dağıtılmış değişik aktivasyon desenlerinin çakışmasıyla ortaya çıktığı düşüncesi ise babaanne nöronu savına rakip bir savdır. Bu sava göre her yüze ait bir nöronun aktivasyonundan daha çok her yüze ait belli bir aktivasyon deseni aktivasyonu oluşması söz konusudur. Değişik yüzlere ait aktivasyon desenleri birbirileriyle ortak ve farklı nöronların aktivasyonlarını içerir. Soyut düzlemlerle birbirinden ayrılabilen bu aktivasyon desenleri birbirilerine ne derece yakınsa kullandıkları nöron devreleri o derece ortaktır ve deseni oluşturan nöron aktivasyonları o derece benzer yüzleri betimlemektedirler. (Dicarlo, 2007, s: 334-335) Bu sav babaanne nöronu savının açıklamakta yetersiz kaldığı bir karikatürden bir insan yüzünü tanınması ya da benzer yüzlerin zaman içinde birbiriyle karıştırılması gibi birçok bulguyu açıklamayı başarır.

Evrişimli sinir ağlarında da söz konusu görsel bilginin sınıflandırılması oluşan aktivasyon desenlerinin lineer bir filtreden geçirilerek yapılacak sınıflandırmaya göre girdinin hangi sınıfa ait olduğunu olasılığı bulunarak yapılmaktadır. Bu tip ağlarda bilginin tek yönde ilerlemesi söz konusudur, dolayısıyla evrişimli sinir ağları ileri beslemeli ağlardır. İleri beslemeli ağlarda bilgi hiyerarşik olarak ilk katmanda işlenmeye başlayarak son katmana doğru aktarılır. Her girdiden sonra bu girdinin hangi sınıfa ait olduğu ağa geri yayılım tekniği kullanılarak öğretilir. İleri beslemeli ağlarda dolayısıyla ağın öğrenmesi ancak çok sayıda etiketli verinin ağa tanıtılması ile mümkün olmaktadır. (LeCun, 2015, s: 436) İleri beslemeli ağları, beyindeki nöronal ağlardan ayıran en önemli fark, bu tip ağların söz konusu görüntülerin temsillerini öğrenmektense söz konusu görüntüdeki düzenlilikleri doğru sınıflandırmayı yapacak şekilde öğrenmeleridir.

### 2.4.3 Tekrarlayan Sinir Ağları ve Gözetimsiz Öğrenme

Beyinde bir nesneyi veya bir yüzü tanımak o nesne veya o yüzle ilgili bilginin hafızadaki temsillerinin nöron hesaplamalarıyla ortaya çıkan bilgiyle eşleştirilmesiyle

mümkün olmaktadır. Bir nesnenin tamamını görmeden de o nesneyi tanımamızda aynı şekilde yapılan hesapların hafızadaki temsillerle kıyasını gerektirir. Görsel illüzyonlardan ve psikolojik çalışmalardan görsel bir algıda beyinde basitten, karmaşığa hiyerarşik bir yapı olmadığı, beklentilerin görsel olarak algılananı etkilediği bilinmektedir. Beyin her zaman için algıları anlamlı kılmaya ve eksiklikleri tamamlamaya meyleder. Psikolojinin Gestalt organizasyon ilkeleri, beynin görsel bir algıyı anlamlandırırken arka planı ihmal ettiğini, benzer nesneleri ilişkilendirdiğini, duyusal verideki eksikleri tamamladığını, aynı anda hareket eden objeleri algısal olarak aynı gruba koyduğunu belirtir. (Goldstein, 2010, s: 105-107) Özellikle Gestalt organizasyon ilkelerinin beynin hiyerarşik işleme yapısıyla açıklanması mümkün değildir.

Beynin duyusal bilgileri işlerken kendi temsillerini kullanarak çıkarım yaptığı fikri ilk olarak Helmholtz tarafından ortaya atılmıştır. Helmholtz retinaya gelen bilginin son derece gürültülü ve kaotik olduğunu, beynin bu karışık veriden gerekli çıkarımları yaparken duyusal algıdaki ipuçlarını beklenti ve inançları doğrultusunda kullandığını belirtmiştir. (Seth, 2021, s: 147) Helmholtz'un beynin duyusal algıdaki ipuçlarını kullanarak beklentileri doğrultusunda dünyayı algıladığı fikrinin bir başka yorumu beynin duyusal bilgiyi yorumlarken en doğru olasılıklı çıkarımı yapmaya çalıştığını iddia eden Bayesci beyin hipotezidir. Bir algı esnasında beyinde bulunan birçok paralel hesaplama bölgesi söz konusu algının bir yorumunu çıkarırken, her aşamada ortaya çıkan yorumun doğru olma olasılığını üst bölgelerden aldığı bilgiyle kıyaslayarak istatistiki olarak hesaplar. Son aşamada ortaya çıkmış olan algı gerçek dünyayla en doğru etkileşimi verecek olan, doğruluk ihtimali en yüksek olan sonuçtur. Dolaysıyla her duyusal algı esnasında beyin sürekli olarak dış dünyada ne olduğuna dair hipotezler üretir. Algımız büyük oranda dış çevreden beyne gelen sinyallerin işlenmesiyle ortaya çıkan bilgiden daha çok her adımda aktif olarak yapılan tahminlere dayanmaktadır. (Seth, 2021, s: 177) Aktif tahminlerin temelinde ise daha önceki deneyimler ile oluşturtulmuş üst düzey temsiller bulunmaktadır. Duyusal algı esnasında bu üst temsillerdeki bilgi daha basit temsillerin oluşturulduğu bölgeleri etkilemektedir. Beynin üst bölgelerinden alt bölgelerine doğru gözlemlenen bilgi akışı "yukarıdan aşağıya işleme" olarak adlandırılmaktadır. Bu süreçte beyin önceden oluşturulmuş olduğu bilgileri anlık algısını şekillendirmek üzere kullanır. Yukarıdan aşağıya işlemede duyusal algının inşası esnasında bu algının ne olabileceğine dair bilgi üst düzey bölgelerden ilk hesapları yapan bölgelere gönderir. (Goldstein, 2010, s: 9-10) Dolayısıyla çevremizi beklentilerimiz ve öğrendiklerimiz doğrultusunda

algılarız. Yukarıdan aşağı işleme hiyerarşik hesaplamalarla açıklanamayan birçok duyusal algıyı ve görsel illüzyonları başarılı şekilde açıklamayı başarmıştır.

Bir başka ifadeyle insan beyni bir görüntünün belli bir nesne olarak tanımlandığını, ileri beslemeli ağlarda olduğu gibi sadece etiketli örnekler üstünden öğrenmez. İlk olarak görsel sinir sistemi çevresindeki nesneleri görmeyi öğrenir, daha sonra gördüğü nesnelerin temsillerini oluşturur. Bir çocuğun bir temsilin kedi olarak adlandırıldığını öğrenmesi ancak kedi temsili sinir sisteminde oluşturulduktan sonra gerçekleşir. Sinir sisteminin kedi temsilini öğrenme süreci ise tamamen kendi başına gerçekleşen bir olaydır. Yapay zeka alanında etikete ihtiyaç duymaksızın ağın belli özellikleri öğrenmesine gözetimsiz öğrenme denmektedir.

#### 2.4.3.1 Hopfield Ağları

Gözetimsiz öğrenmenin temelleri aynı zamanda ilk tekrarlayan sinir ağı örneği olan Hopfield ağlarına dayanır. Tekrarlayan sinir ağlarının evrişimli yapay sinir ağları ve genel olarak ileri beslemeli yapay sinir ağlarından farkı nöronlar arasındaki bağlantıların çift yönlü olmasıdır. İleri beslemeli ağlarda işlenecek olan bilgi ilk girdi katmanından itibaren ağda sadece ileriye doğru tek yönde ilerler. Ancak tekrarlayan sinir ağlarında bilgi girdi katmanından başlayarak çift yönlü dinamik bir süreçte ağın yapısına nüfuz eder.

Hopfield ağları, 1982 yılında Hopfield tarafından tasarlanmıştır. Fizikteki Ising modellerinden yola çıkan Hopfield, enerji tabanlı bir ağ kullanarak ilişkisel hafızanın yapay sinir ağlarınca oluşturulabileceğini göstermiştir. Hopfield ağlarının tasarlanma amaçları belli bir örüntünün öğrenilmesinden daha çok öğrenilmiş bir örüntünün nasıl ilişkisel hafızada muhafaza edileceğinin gösterilmesi olmuştur. Hopfield ağı simetrik bağlantılarla birbirine bağlanmış açık veya kapalı olmak üzere iki değer alabilen nöronlardan oluşur. Her nöron arasındaki bağlantı simetriktir ve nöronlar sadece lokal bağlantılara sahiptir, dolayısıyla bir nöron sadece komşu olduğu nöronlarla bağlantılıdır. Hopfield, fizikten ilham alarak nöronların birbiriyle etkileşimlerinin bir enerji fonksiyonu aracılığıyla tanımlamış ve fiziğin bir kaidesi olarak ağın stabil olduğu durumlarda ağın toplam enerjisinin minimum enerji seviyesinde olacağını belirtmiştir. (Hopfield, 1982, s: 2554) Ancak nöronlar birbirilerine lokal olarak bağlı olduğundan bu tip bir ağ minimum enerji seviyesine ancak iki boyutlu bir uzayda lokal minimum enerji noktaları oluşturarak ulaşma imkanına sahiptir. Hopfield ağlarının belli bir zamanda

toplam enerjilerini minimize etmek için lokal minimumlar oluşturması bu ağların çift yönlü dinamik yapısından kaynaklanır. Ağa hatırlaması gereken bir bilgi yerleştirildiğinde ağ, bu bilgiyi saklayabileceği en düşük toplam enerjili konfigürasyonu aramaya başlar ve bu arayış çift yönlü lokal etkileşimlerle mümkün olur. Ağdaki tüm nöronlar denge konfigürasyonuna ulaşana dek ağdaki nöronların açık ya da kapalı olma konumları komşu nöronların durumuna göre karşılıklı etkileşimlerle değişir. Ağ denge noktasına ulaşıldığında ağda lokal olarak belli sayıda nöron tarafından oluşturulan minimum enerjili konfigürasyonlar ağa verilen bilginin bir temsilini oluşturur. (Hopfield, 1982, s: 2555-2257)

Bu tip enerji tabanlı ağların en önemli özellikleri verilen bilginin dinamik temsilerini oluşturmalarıdır ki bir ağın dinamik temsiller oluşturabilme yetisi bu ağlara ileri beslemeli ağlara göre üstünlük sağlar. Hopfield ağının ilişkisel hafıza oluşturması da bu tip üstünlüklerden biridir. Hafızasını oluşturmuş bir Hopfield ağına yeni bir bilgiyi verildiğinde ağın yeni bilgideki eksiklikleri tamamlaması veya bilgideki belli bir hatayı içsel temsiliyle kıyaslayarak düzeltebilmesi ileri beslemeli ağların sahip olmadığı özelliklerdir. Bir Hopfield ağlarının oluşturduğu hafıza birimleri verinin içeriğine göre belli bir noktada saklanmasıysa bu ağları verinin içeriğine göre adreslemiş bir hafıza sistemine dönüştürür. Hopfield ağlarının verideki eksikleri tamamlama, yanlış verileri kendi temsilleri ile düzeltme ve veriler arasında ilişki kurabilme gibi özellikleri bu ağların insan hafızası ile ortak noktalarıdır. (Hopfield, 1982, s: 2558)

Hopfield ağlarının güçlü taraflarının yanı sıra oldukça problemli tarafları da bulunmaktaydı. Bu ağların en büyük dezavantajı kullanılan nöron sayısıyla kıyaslandığında hafızalarında saklayabildikleri veri miktarının çok düşük olmasıydı. Ayrıca Hopfield ağlarında eğer iki düşük enerji noktası ağda birbirine yakın alanlarda oluşursa ağın dinamik yapısı bu iki düşük enerjili konfigürasyonu birleştirerek sahte anılar ortaya çıkarıyordu. Hopfield ağlarındaki bu sahte anılardan kurtulmak için bulunan bir yöntem ise uyanıklık ve uyku algoritmasıydı. Hopfield, bu algoritmayı, Crick'in, rüya görmenin işlevinin gereksiz bilgilerin unutulması olduğu iddiası üstüne kurmuştur. (Hopfield, 1983, s: 158-159) İlk faz olan uyanıklık fazında ağa dış veriler yükleniyor ve ağın kendi enerjisini minimize ederek dengeli bir duruma geçmesi yani hafızasını oluşturması bekleniyordu. İkinci faz olan uyku fazında ise dış veri olmaksızın ağ aktive ediliyor ve ağın yeni denge halindeki ulaştığı minimum durumlar ağdan siliniyor, yani ağın bu durumları unutması sağlanıyordu.

Unutmanın öğrenmenin gerçekleşmesi için elzem olduğu, Hebb kanunun doğru ancak eksik olduğu ise yakın zamanda anlaşılmıştır. İnsan beyninde iki tip hafıza birimin varlığından söz etmek mümkündür. Kısa dönemli hafıza, seçici dikkat mekanizmalarıyla yakından alakalıdır ve dikkatin yönlendirildiği belli miktarda bilgiyi kısa bir süre boyunca hatırlamamızı sağlar. Kısa dönemli hafızanın limiti kısıtlıdır ve kısa dönemli hafızada ortalama yedi parça bilgi saklanabilir. Kısa dönemli hafıza aynı zamanda anlık algılarımızın devamlılığını sağlayan bir süreçtir. Kısa dönemli hafızadaki bilgiler, hipokampüsün işleviyle kısa dönemli hafızaların bir toplamı olarak tanımlanabilen tek bir uzun süreli hafızaya çevrilir. Kısa dönemli hafıza ile uzun süreli hafıza arasında tanımlanan çalışan hafıza kısa dönemli hafızadaki bilgilerin aktif olarak manipüle edildiği geçiş mekanizmasıdır. Uzun süreli hafıza kapasitesi limitsizdir ve bilgileri çok uzun süre depolama yetisine sahiptir. (Luo, 2016, s: 417)

Hebb öğrenme kanunun açıklamakta yetersiz kaldığı en temel nokta organizmanın her duyusal girdiyi aynı şekilde öğrenmiyor oluşudur. Art arda gelen iki bilgiden ilki organizma için önemsiz, ikincisi önemli olduğunda ilk bilginin bastırıldığı, dolayısıyla unutulduğu gözlenmiştir. Ancak ilk bilgi organizma için önemsiz, ikinci bilgi organizma için önemliyse, iki bilgide hafızada yer etmekte ve birbirleriyle ilişkilendirilmektedir. 1973 yılında Bliss ve Lomo, bir tavşanın prefontal yolunu stimule etmiş ve bölgedeki sinaptik aktivitenin hatırı sayılır bir süre boyunca arttığını gözlemlemiş, glumat nöro ileticisiyle oluşan bu aktivitenin beyindeki fiziksel değişiklikleri tetiklediğini görmüştür. Uzun süreli potansiyasyon olarak adlandırılan bu aktivitenin uzun süreli belleğin oluşumunda ilk adım olduğu anlaşılmıştır. Uzun süreli potansiyasyon mekanizmasının ters mekanizmasının ise uzun süreli depresyoun mekanizması olduğu keşfedilmiştir. Uzun süreli bastırma mekanizması aynen uzun süreli potansiyasyon mekanizması gibi hücre seviyesinde gerçekleşen kimyasal bir süreçtir. Uzun süreli depresyonda sinapstik aktivitenin hatırı sayılır bir süre boyunca bastırılması sonucu yeni bilginin öğrenilmez veya hemen unutulur. Uzun süreli potasiyasyon ve depresyon mekanizmalarının ağırlıkla gözlemlendiği bölgeler hipokampal formasyona ait yapılardır. (Luo, 2016, s: 421-434) Hebb öğrenme kanununda uzun süreli depresyon mekanizması bulunmamaktadır. Uyku uyanıklık algoritmaları uzun süreli depresyon mekanizmasını taklit ederek iş görürler.

### 2.4.3.2 Boltzman makineleri ve İnanç ağları

Hinton, Hopfield ağlarının hafıza birimi olarak kullanmaktan daha çok veriyi öğrenebilen ve öğrendikleri verileri yeni verilerle ilişkilendirilerek yorumlayabilen ağlar olarak tasarlamanın olanaklı olup olmadığını araştırırken Boltzman makinelerini tasarlamıştır. Boltzman makineleri görünür ve saklı katmanlara sahip çok katmanlı Hopfield ağlarıdır. Boltzman makineleri de çift taraflı simetrik ağırlıklarla birbirine bağlanmış nöronlardan oluşlar. Hinton, Boltzman makinelerinin görünür katmanını girdi verisini temsil edecek şekilde tasarlamış, böylece öğrenebilen bir sinir ağı elde etmeye çalışmıştır. (Hinton, 2014, s: 1-3) Hopfield ağları ilişkisel hafıza birimleri olarak iş görürken, Boltzman makinelerinin belli verileri öğrenmesi ve daha sonra öğrendikleri verilerden yola çıkarak herhangi bir veriyi yorumlaması istenmekteydi. Bir veriyi, daha önceki verileri göz önünde bulundurarak yorumlamak ise yeni verinin tüm eski verilerle benzerliklerini bir olasılık dağılımında gösterilmesini gerektiriyordu. Bu nedenle Boltzman makineleri, Hopfield ağlarındaki enerji modelleriyle iş görmekten daha çok termodinamik yasalarıyla iş görecek şekilde tasarlanmıştır. Dolayısıyla Hopfield ağındaki ikili durumlu nöronlar yerine rastgele davranan stokastik ikili durumlu nöronlar kullanılmış ve bu nöronların enerji fonksiyonları ısıya bağlı olarak tanımlanmıştır. Bu şekilde tasarlanan ağ termal dengeye ulaştığında değişik enerji seviyelerinin dağılımı sabit bir olasılık dağılımı yani Boltzman dağılımı gösterir. Ancak Boltzman makinesinin termal dengeye ulaşması sistemin en az enerjili duruma ulaşmasıyla aynı anlamı taşımaz. Ağ termal dengeye ulaştığında gruplaşmış nöron kümelerinin aktivitesi devam eder ancak bu aktivite belirlenmiş bir olasılık dağılımına sahip olur. Yapılan çalışmalarda Boltzman makinelerinin eğitilmesinin oldukça kaotik ve zor olduğu anlaşılması üstüne Boltzman makinelerinin tasarımı sadeleştirilerek kısıtlanmış Boltzman makinelerini tasarlanmıştır. (Hinton, 2014, s: 5)

Kısıtlanmış Boltzman makinelerinde girdinin bağlandığı görünen nöronlar arasında bağlantı bulunmadığı gibi saklı katman nöronları arasındaki bağlantılarda iptal edilmiştir. Genel olarak görünür ve gizli katman olarak iki katmanlı olarak tasarlanan Kısıtlanmış Boltzman makinelerinin tasarımındaki sadeleştirme bu ağların gizli katmanlarındaki olasılık dağılımı hesaplanabilir kılmıştır. Kısıtlanmış Boltzman makinelerinde hedef verilen girdinin en verimli temsilini iki boyutlu bir düzlem oluşturan gizli katmanlarında tekrar girdi verisini üretebilecekleri şekilde oluşturabilmektir. Kısıtlanmış Boltzman makinelerinin eğitimi için sadece girdi verisi yeterlidir, en nihayetinde Kısıtlanmış Boltzman makineleri girdi verisinin tekrar

oluşturabilecek en verimli temsili gizli katmanlarında oluşturmak, bir başka ifade ile girdi verisinin içindeki örüntü ve bağlantıları öğrenerek girdi verisini en verimli şekilde modellemek için eğitilirler. Bu eğitimde uyku ve uyanıklık algoritması kullanılmaktadır. Uyanıklık fazında ağın temsil etmesi istenen girdi ağa verilmekte ve uyku fazında ağın bu girdiyi en iyi şekilde yeniden üretmesi için gerekli düzenlemeler yapılmaktadır. (Hinton, 1995, s: 1158-1162)

Dış duyusal veriyi kendi temsillerinde modelleme yetisi olan tüm yapay sinir ağları aynı zamanda "üretken sinir ağları" olarak tanımlanırlar. Bu ağlar dış bir girdi olmaksızın daha önce oluşturdukları temsillerden kendi çıktılarını da oluşturabilirler ve ileri beslemeli ağlarla kıyas kabul etmez bir esnekliğe sahiptirler. Örneğin bir Boltzman makinesi el yazısı ile yazılmış iki rakamını okumak için eğitildiğinde ağ, kendi iç temsillerinde her nöron grubunun iki rakamın belli özelliklerini belli nöron desenleriyle temsil eder. Bu şekilde eğitilmiş bir Boltzman makinesi simetrik bağlantılara sahip olduğundan sadece iki rakamını tanımakla kalmaz, aynı zamanda kendi el yazısı iki rakamını üretebilir. İki rakamının temsilini oluşturmuş bir Boltzman makinesine üç rakamı gösterildiğinde ise Boltzman makinesi girilen üç rakamını iki rakamına benzetmeye çalışır. Ağın iç temsilinde iki rakamının temsili olduğundan ağ verilen her girdiyi iki olarak görme eğilimindedir ancak bu eğilim aynı zamanda ağın başarısının temel noktasıdır.

Kısıtlanmış Boltzman makinelerinin art arda bağlanması ile "derin sinir ağları" elde edilir. Örneğin görüntü tanıyan derin sinir ağlarının tasarımında piksellerden girdi alacak olan ilk katmandaki kısıtlanmış Boltzman makinesi görüntüdeki herhangi bir lineer olmayan özelliği öğrenecek şekilde eğitilir. Daha sonra ilk kısıtlanmış Boltzman makinesinin çıktısı ile ikinci katmandaki kısıtlanmış Boltzman makinesinin girdisi oluşturur. Böylece belli özellikleri öğrenmiş olan ilk Boltzman makinesiyle beraber ikinci Boltzman makinesi de girdi pikselleri yeniden üretecek şekilde eğitilirler. Bu eğitim istenilen derinlikteki sinir ağı eğitilene dek devam eder. Pratikte Kısıtlanmış Boltzman makineleri ikili katmanlar olarak eğitilirler, sadece iki Boltzman makinesinin birbirilerinin girdi ve çıktılarını üretecek şekilde eğitilmesine açgözlü öğrenme denmektedir. (Bengio, 2007, s: 1-2) Sınırlı Boltzman makinelerinin ikili çiftler olarak eğitilmesi hızlı bir yöntem olduğu için tercih edilmektedir. Birçok Boltzman makinesinin kullanıldığı derin ağlarda öğrenmenin sadece komşu katmanların etkileşimi ile yapılması oldukça verimli sonuçlar vermektedir.

Derin ağın eğitimi tamamlandığında her katmanda en basit özelliklerin temsilleri ile başlayıp, gitgide karmaşıklaşan özelliklerin temsili oluşur. Ağın her katmanda en basitten başlayarak gitgide daha komplike özellikleri temsil etme özelliği ağa el ile tasarlanarak verilmiş bir özellik değildir. Boltzman makineleri derin sinir ağındaki hiyerarşik temsilleri sadece kendi iç dinamikleriyle oluştururlar. Derin sinir ağlarının oluşturmuş olduğu bu hiyerarşik organizasyon beynin duyusal sistemlerinde gözlemlenen organizasyonuyla aynı paralelliğe sahiptir. Örneğin görmek için eğitilmiş bir derin sinir ağının oluşturduğu hiyerarşik yapıda ilk katmanlar birincil görsel alanın işlevlerini yapmaya başlar. (Yamins, 2016, s: 361) Eğitilmiş kısıtlanmış Boltzman makinelerinden oluşan derin ağ, piksellerin temsillerini en alt seviyeden en üst seviyeye dek oluşturmuştur. Derin sinir ağının ince ayarı ise derin ağın girdideki veriyi kendi temsillerinde en doğru şekilde oluşturması için yapılır. Bu aşamada ağ zaten dış duyusal veriyi işlemeyi öğrenmiş olduğundan az sayıda örnekle ağ kendi temsillerini girdi temsillerini tekrar oluşturabilecek şekilde oluşturur. Tekrarlayan sinir ağlarının sınıflandırma problemlerinde dahi ileri beslemeli ağlardan daha üstün olmalarının nedeni ileri beslemeli ağların sadece dünyaya ait görseldeki piksellere dayanarak sınıflandırırken tekrarlayan sinir ağlarının direkt olarak dünyaya ait görsel temsilleri öğrenmeleridir. Görmek için eğitilmiş bir derin sinir ağı iki değişik örüntüyü tanıması için eğitilirse, eğitimin belli bir noktasında ağ iki farklı örüntüyü iki farklı kategori gibi algılamaya başladığı ve hafızasını ikiye ayırdığı gözlemlenir. (Yamins, 2016, s: 356) Ağın bu özelliği insan beyninde örneğin yüzleri ve nesneleri tanıyan bölgelerin işlevsel olarak ayrılmasıyla kıyaslanabilir. Derin sinir ağlarının insan beyninin organizasyonu ile paylaştığı bir başka noktaysa eğitimi tamamlanmış bir derin sinir ağından belli bir temsili üretmesi istendiğinde alt katmanlardan daha çok temsili en üst katmanların üretmesidir. (Yamins, 2016, s: 359-360)

Yineleyen bir başka yapay sinir ağı örneği ise 1992 yılında B. Neal tarafından tasarlanan sigmoid inanç ağlarıdır. Sigmoid inanç ağları beynin duyusal verileri çift yönlü bir iletişimle işlemesinden esinlenerek oluşturulmuş yapay sinir ağlarıdırlar. İnanç ağlarında da stotastik nöronlar kullanılmıştır ve bu nöronlar arasındaki bağlantılar çift taraflıdır. İnanç ağlarında hedeflenen amaç girdi verisinin yapısını ağın iç temsillerinde modellenmesi ve bu modellerdeki olası dağılımların girdi verisinin temel özelliklerini yeniden üretebilmesidir. Ancak sigmoid inanç ağlarının çift yönlü dinamik yapısı bu ağların eğitilmesindeki en büyük zorluk olmuştur. Yapay sinir ağının eğitilmesi ağdaki nöronlar arasındaki ağırlıkların ayarlanması olarak ifade edilirse inanç ağındaki bir nöronun komşu nöronlarla arasındaki ağırlığın ayarlanması,

nöronun hem önceki katmandaki nöronlar, hem de sonraki katmandaki nöronlara karşılıklı etkileşiminin dinamiğine bağlıdır. İnanç ağları, enerji tabanlı ağlar gibi lokal öğrenme kurallarına sahiptir. Bir inanç ağına yeni bir veri verildiğinde bu verinin ağın o ana dek oluşturduğu temsili de değiştirecektir. Dolayısıyla inanç ağlarının öğrenmesi sırasında nöronlar arasındaki ağırlıkların iki farklı etki altında olduğu düşünülebilir. Ağın verideki bilgiyi aşağıdan yukarı taşıyan bağlantıları nöronlar arasındaki ağırlıkları girdi verisiyle mümkün olduğunca aynı bir iç temsil oluşturacak şekilde değiştirmeye çalışırken, yukarıdan aşağı bağlantılar girdi verisini ağın o ana dek oluşturduğu temsille uyacak şekilde değiştirmeye çalışacaklardır. Dolayısıyla ağın öğrenmesi bu iki etkinin arasındaki farkın en aza indirilmesi anlamını taşımaktadır. Bir başka deyişle bu tip ağlarda öğrenmenin amacı ağa gösterilen gerçek temsil ile ağın yaratmış olduğu temsilin mümkün olduğunca birbirileriyle uyumlu hale getirilmesidir. (Neal, 1992, s: 71-113)

Uzun bir süre boyunca Sigmoid inanç ağları ile kısıtlanmış Boltzman makineleri değişik temelde çalışan ağlar olduğu düşünülmüştür. Hopfield ağları ve Kısıtlanmış Boltzman makineleri enerji tabanlı ağlar iken sigmoid inanç ağları nedensellik temelli ağlar olarak tasarlanmışlardır. Ancak üst üstte bağlanan kısıtlanmış Boltzman makinelerinden oluşturulan derin sinir ağlarının sigmoid inanç ağlarına eşdeğer olduğunun anlaşılması sigmoid inanç ağlarının enerji tabanlı ağlar olarak tanımlanmalarının mümkün olduğunu göstermiştir. (Hinton, 2008, s: 2630-2635)

Bayesci beyin modelleri, beynin iç temsillerini kullanarak duyusal algılardaki ipuçlarını en doğru çıkarımı yapacak şekilde yorumladığını iddia etmiştirler. Ancak bu modeller beynin söz konusu çıkarımları nasıl yaptığını açıklamamıştırlar. Friston, Bayesci modellerin nasıl iş gördüğünü “serbest enerji prensibi” ile fiziksel bir temele oturtmuştur. Serbest enerji prensibi tek bir hücreden insan beyni gibi en karışık biyolojik sistemlere uygulanabilen bir prensiptir ve kendi kendini organize eden tüm biyolojik yapıların kendi bütünlüklerini korumak için dış dünya ile dış dünyaya dair kendi iç temsilleri arasındaki ön görü farkını en aza indirmeye çalışır gibi bir davranış sergilediklerini iddia eder. Dolayısıyla çevrenin özellikleri kendi kendini organize eden sistemler için hayati öneme sahiptir. Beyin duyusal algıdan doğru çıkarımlar yapabilmek için ilk olarak iç temsillerinde dış dünyanın bir modelini üretmek zorundadır. Beynin dış dünyayı kendi iç temsillerinde ürettiğinde, her yeni duyusal algıyı önceden oluşturmuş olduğu iç temsilleri kıyaslayarak, iç temsillerinin dış dünyayı doğru ürettiğine dair kanıtlar aramaktadır. Fiziksel olarak bu durum beynin

dış duyusal bilgi ile iç temsilleri arasındaki farkı, yani serbest enerjiyi minimize etmeye çalışması demektir. (Friston, 2012, s: 1232-1233) Dolayısıyla Friston'a göre sigmoid inanç ağları, beynin bir modelidir.

Tekrarlayan sinir ağlarının dış dünyanın temsillerini içsel modellerinde yaratmalarının en çarpıcı özelliği beynin bilinçli içeriğinin yukarıdan aşağı işlemeyle alakalı bir kavram olduğu düşüncesinden kaynaklanmaktadır. Beynimiz dış dünyayı içsel temsilleri aracılığıyla kendi yapısında oluşturur, beynimizdeki temsiller dış dünyanın özelliklerini içerirler. Temsilciliğe göre bir şeyi bilinçli olarak deneyimlememiz, o şeyin zihinsel temsili ile ilgilidir. Tüm zihinsel temsiller bilinçli değildir ancak bilinçli her deneyim beynin içsel temsillerini içerir. Eğer ki beynimizde yukarıdan aşağı bir işleme söz konusu ise, yukarıdan aşağı işlemeyi meydana getiren zihinsel temsilin bilinçli bir içerik oluşturması muhtemeldir. Söz konusu zihinsel temsillerin hangilerinin ve hangi durumlarda bilinçli içerik oluşturduğu sorusu ise bilinçli içeriğin nasıl meydana geldiği sorusuyla yakından alakalıdır. Ancak bilinçli içerikler her zaman için ikinci dereceden bir temsildir; dış dünyanın yaratılmış içsel temsilinin tekrardan temsilidirler.

## 2.5 NÖROBİLİMİN BİLİNÇ MODELLERİ

Bilim dünyasında bilinç çalışmalarının yaygınlaşması nispeten oldukça yeni bir gelişmedir. Üçüncü şahıs gözlemlerine dayanan bilimsel yöntem nedeniyle bilişsel bilimler birinci şahsın deneyimlerine dayanan bilinci incelemekten uzun süre kaçınmıştır. Francis Crick, 1990 yılında kaleme almış olduğu *Şaşırtan Varsayım* kitabında nörobilim dalında çalışan bilim insanlarının bilinci merak etmekte son derece isteksiz olduklarını vurgular. Crick, bilimin bilinci araştırmaktan kaçınmasının nedenleri arasında davranışçı ekolün bilinci içine girilmez bir kara kutu olarak göstermiş olmasının yanı sıra bilinci incelemek adına doğru bilimsel yöntemlerin belirlenmemiş olması olduğunu belirtmiştir.

Crick, bilincin bilim tarafından çalışılabilir olduğunu ve bilimin indirgemeci yöntemleriyle bilinç çalışmalarından sonuç alınamayacağını belirten tüm felsefi itirazlara rağmen bilinci çalışma yönteminin indirgemeci yöntem olması gerektiğini savunmuştur. İndirgemecilik fizik, kimya ve moleküler biyolojinin gelişmesini hızlandıran başlıca kuramsal yöntem olmuştur. Modern bilimdeki baş döndürücü ilerlemenin altında yatan, indirgemeci yöntemdir. (Crick, 1997, s: 5) Crick, bilincin

tamamen beynin çalışma prensiplerinin bir ürünü olduğunu belirtir. Crick bu varsayımı "şaşırtan varsayım" olarak adlandırır. "Şaşırtan varsayım: siz, neşeleriniz, üzüntüleriniz, anılarınız, ihtiraslarınız, benlik ve özgür irade duygularınızla, aslında çok sayıda nöron ve bunlarla ilişkili moleküllerin bir arada davranışından ibaretsiniz." (Crick, 1997, s: 1) Crick'in şaşırtan varsayımı, tüm bilimsel araştırmalardaki bilinç çalışmalarının temelindeki varsayımdır.

Crick'in ikinci varsayımı ise acı veya mutluluk hissetmek, görsel veya duyumsal farkındalık gibi bilincin çeşitli yönlerinin ortak bir işleyişin ya da düzenin sonucu olduğudur. Dolayısıyla Crick, bilinçli bir deneyimin, örneğin bir görselin bilinçli biçimde idrak edilmesinin nasıl gerçekleştiğini ya da beyinde görsel bilinçli deneyimin hangi nöronal sistemlerin hangi dinamikleriyle ortaya çıktığını bulduğu taktirde genel olarak tüm bilinçli deneylerimin açıklanabileceğini savunmuştur. (Crick, 1997, s: 23)

İki temel varsayıma ilave olarak Crick, bilinç için bilimsel bir araştırma zemini belirlerken bilincin kesin bir tanımını yapmaktan kaçınmıştır. Crick yapılacak bir tanımın bilinç araştırmalarını kısıtlama ve yanlış bir yola yönlendirme olasılığı olacağını belirtmiştir. Crick ve Koch, bilinci bilimsel bir metotla çalışmanın en basit yolunun beyinde bilinçli deneyime neden olan minimal nöron bağlantılarının bulunması olduğunu ve bilinçli deneyimin minimal nöronal bağlantılarının görsel sistemdeki bilinçli deneyimi inceleyerek bulunabileceğini öne sürmüşlerdir. (Crick, Koch, 1990, s: 266-267) Görsel sistem korteksin üstüne en ayrıntılı çalışılmış bölümüdür ve görsel illüzyonlar bilinci deneysel metotlarla çalışmak için oldukça zengin bir repertuar sunar. Crick ve Koch, görsel illüzyonlarda var olan bir şeklin bilinçli deneyim yarattığı ve yaratmadığı durumlarda beyin aktivitesini inceleyerek hangi nöronların ya da nöronların hangi etkileşimlerinin bilinçli deneyime neden olduğunun bulunmasının bilinç araştırmalarındaki ilk adım olduğunu savunur. Bir başka araştırma yolu ise beynin bir görselden değişik bilinçli deneyimler çıkardığı durumlardır. Bu duruma "nektar küpü" örnek verilebilir. Bir nektar küpünü iki değişik bilinçli deneyim yaratır, değişik anlarda küpün iki olası ön yüzünden sadece biri ön yüz olarak görünür. Eğer nektar küpünün bilinçli görsel deneyimde içeriği değiştiğinde beyinde hangi nöron aktivite ve bağlantılarının değiştiği gözlenebilirse, bu gözlem bilinç içeriği ile nöronal aktivite arasındaki bağlantıyı ortaya çıkaracaktır. (Crick, Koch, 1990, s: 272-273)

Beyin korteksinde çoğu aktivitenin bilinçli deneyim yaratmaz. Bu varsayımı destekleyen uç bir örnek "kör görüşü"dür. Kör görüşü birincil görme korteksinde

meydana gelen bir hasar veya aynı bölgede oluşan bir tümörle bilinçli görme yetisini kaybeden hastalarda karşılaşılan bir durumdur. Kör görüşüne sahip hastalar, kör olduklarını dolayısıyla göremediklerini beyan ederler. Ancak bu hastalardan rastgele konmuş değişik objelerin olduğu bir koridorda yürümeleri istendiğinde, hastalar hiçbir objeye çarpmadan koridorun sonuna ulaşmayı başarır. Bu hastalar bilinçli olarak görmedikleri halde beyin görsel bilgiyi işlemekte ve hastanın hareketini düzenlemektedir. Kör görüşüne sahip hastalar, objelere çarpmadan yollarını bulmanın ötesinde yetenekler de sergilerler ancak bunların hiçbirinin bilincine varamazlar. Birincil görme korteksi çıkarılan makak maymunlarında yapılan deneylerde bilinçli görme meydana gelmeksizin, beyindeki sistemlerin görsel bilgiyi belli bir oranda işleyebildiğini göstermiştir. (Koch, 2004, s: 222)

Beynimizde bilinçli deneyim yaratmaksızın çalışan bu tip sistemlere “zombi sistemler” denmektedir. (Koch, 2004, s: 220) Zombi sistemler, beyin sapı veya orta beyin yapıları tarafından oluşturulan refleks sistemlerini kapsamakla beraber, reflekslerden farklıdır. Zombi sistemler, rutin aktiviteleri bilinçli deneyim yaratmaksızın gerçekleştiren sistemlerdir. Beyin korteksi aktivitesinin sadece belli alt kümeleri bilinçli deneyimle ilişkilendirilebilir gibi görünmektedir. Beynimiz birçok işlevini bilince yansımadan yerine getirir. Beyin korteksinde, bilinçli içeriklerle alakalı olup direkt olarak bilinçli deneyim yaratmayan birçok bölge bulunmaktadır. Konuşurken, konuşmanın içeriğine dair bir fikre sahibizdir ancak seçtiğimiz kelimeler, kurduğumuz cümleler, üstüne düşünmeden ağzımızdan çıkar. Kelimelerden belli dilbilgisi kurallarıyla cümleler oluştururken veya konuşurken dilimizi ve dudaklarımızı oldukça komplike şekilde hareket ettirirken ne yaptığımızın ve bu işlevleri nasıl yerine getirdiğimizin farkında değilizdir. Yeni bir beceriyi öğrenirken yaptığımız her şeyi kontrol etme ihtiyacı duyarız, fakat belli bir süre sonra performans otomatikleşir, ardından performansın birçok adımı bilinçten silinir ve bu aşamada söz konusu performansı yerine getirmek zombi sistemler tarafından üstlenilir. Bilişsel yaşamımızın çoğu otomatikleşmiş rutinlerin ürünüdür. Konuşmayı öğrenmek için uzun zaman bilinçli bir çaba göstermişizdir, fakat sonrasında bu süreç otomatik hale gelmiştir. (Edelman, Tononi, 2019, s: 84)

Bilincin nöronal bağlantılarını tespit etmekteki en büyük zorluk ise bilinçli deneyimi yaratan minimal nöronal sistemleri beynin aktivitesinin geri kalanından ayırmaktır. Bir görselin algılanması sırasında görsel sistemin tamamında, dikkat mekanizmalarını kapsayan bölgelerde, bilinçli deneyimi rapor eden ve değerlendiren frontal bölgede,

kısacası beyindeki birçok değişik modülde birçok nöronal aktivite gerçekleşmektedir, ancak bu aktivitelerin küçük bir kısmı bilinçli deneyimin içeriğini oluşturmaktadır. Ayrıca bilinçli olmakla alakalı olup, bilinçli deneyime katkı sağlamayan beyin yapıları vardır. Beyin sapının işlevinin yitirmesinin bilinçli deneyimleri olanaksız kıldığı, tartışmasız kabul edilen bir durumdur. Crick ve Koch'a göre, beyin sapı ve beyin sapı çevresindeki çekirdeklerin aktivitesi bilincin oluşabilmesi için gerekli arka plan koşuludur ancak beyin sapı bilinçli deneyim yaratmaz. Beyin sapı gibi bilinç için elzem olan ama bilinçli deneyimin yaratılması ve bilinçli deneyimin içeriği ile alakası olmayan bu tip sistemler, bilincin var olması için gerekli arka plan koşulları olarak adlandırılır. (Tonini, 2016, s: 452) Crick ve Koch'a göre bu koşullar bilinçli deneyimi yaratan mekanizmalardan ve bilinçli deneyimin içeriğinin oluşturulma süreçlerden titizlikle ayırt edilmelidir.

Bilinçli deneyimin bir başka özelliği ise çok çeşitli içerikleri kapsamakla beraber her zaman bir bütün oluşturmasıdır. Bilinç her zaman tek bir düşünce ile uğraşır. (Deheane, 2018, s: 215) Bilincin bütünlüğünü vurgulayan en önemli deneylerden biri "iki gözün rekabeti deneyi"dir. İki göze farklı iki görüntü gösterildiğinde, örneğin bir göze yatay çizgiler, bir göze dikey çizgiler gösterildiğinde, beyin bu görüntüleri üst üstte bindiremez, iki görselden birini görür ve görünen görüntü belli bir frekansla değişir. Bilinçli deneyimin her zaman tek ve bir bütün yaratacak şekilde birleştirilmiş bir içeriği vardır. (Deheane, 2018, s: 45)

Bilinçli bir deneyimin içeriğinden bahsedilebileceği gibi bilincin seviyelerinden bahsetmek de mümkündür. Nörobilim bilincin yoğunluk derecelerini ölçmek adına değişik görüntüleme tekniklerinden faydalanır. Beyin ölümü gerçekleşmediği sürece beyin korteksinde kesintisiz nöronal bir etkinlik mevcuttur. Nöronların ateşlenmeleri zamana bağlı fizyolojik bir süreçtir ve beyin aktivitesini gözlemlemenin bir yolu ateşlenen nöronların ortalama ateşlenme frekansının ölçülmesidir. Beyindeki nöron topluluklarının ateşlenmesi ritmik bir örüntüye sahiptir. EEG beyindeki ritmik aktiviteyi zamana bağlı olarak kaydedebilen bir teknolojidir. Beyindeki nöron topluluklarının ritmik aktivitesine, genel olarak beyin dalgaları denmektedir. Sağlıklı bir beyinde dört tip beyin dalgası gözlemlenir ve söz konusu beyin dalgaları bilinç durumuna göre değişiklik gösterir. Bu teknolojiden faydalanarak yapılan ölçümlerde, en yüksek bilinçli durum olan tetikte olma durumunda beyinde 12 hertz ile 30 hertz arasında değişen Beta dalgaları baskınken, uyanık ama tetikte olmayan bir beyinde 8 Hertz ile 13 Hertz aralığında değişen Alpha dalgaları gözlemlenir. Hafif uykuda ise beyinde 4 Hertz ile

7 Hertz aralığındaki Tetha dalgaları baskındır. Uyku süreci ise derin uyku ve REM uykusunu içerir; derin uyku bilinçsiz uyku olarak adlandırılır ve beyinde ortalama 2 Hertzlik Delta dalgaları baskındır. REM uykusunda ise beyin dalgalarının değişik skalalarda değişmekte olduğu gözlenmiştir. Derin uykuda bilinçten söz edilemezken, rüya görmenin gerçekleştiği REM uykusunda bilinçli bir durum söz konusudur. (Tonini, 2016, s: 458-459)

Bilincin yoğunluk derecesini değiştiren diğer durumlar ise anestezi, bitkisel durum ve koma olarak tanımlanmaktadır. Komada beyin korteksinde anlamlı hiçbir aktivite gözlemlenmezken, bitkisel durumda olan bir hastada beynin biyolojik saatine bağlı fonksiyonlar çalışmaya devam eder, ancak hasta minimal bir bilinç durumundadır. Anestezi altında ise beyinde hatırı sayılır bir aktivite varken bilincin varlığı söz konusu değildir. Locked-in sendromundaysa hasta belli bir ölçüde bilinçlidir ancak motor hareketleri olmadığından dış dünya ile iletişim kurması mümkün değildir. Yapılan çalışmalarda bilinçli durumları bilinçsiz durumlardan ayırt etmenin kaba bir yolu da beynin bir bölgesini dışarıdan bir elektrot aracılığıyla aktive ederek aktivitenin beyindeki yayılımını izlemektir. Bilinçli bir beyinde aktivasyon tüm beyne yayılırken, bilinçsiz durumlarda aktivitenin lokalize olduğu, diğer beyin bölgelerine yayılmadığı ve bilinçli bir duruma kıyasen aktivitenin hızla son bulduğu gözlemlenmiştir. (Edelman, Tonini, 2019, s: 83) Bilincin değişik yoğunluk dereceleri, beyin korteksindeki dağınık ve aynı zamanda bütünleşik aktiviteyle doğru orantılı gözükmektedir. Tıp henüz koma halinde gözüken bir hastanın ne kadar bilinçli olduğunu ya da yeniden bilinçli hale gelip gelemeyeceğini öngörememekle beraber son yıllarda beynin bütünlüklü aktivitesi ve bilinçli deneyim arasında kurulan bağlantı sayesinde bilinç kaybı yaşayan hastaların tedavisinde daha doğru öngörülerde bulunulabilmektedir.

### 2.5.1 Kapsamlı Çalışma Alanı Modeli

Günümüzde hala geçerliliğini koruyan bir bilinç modeli ilk olarak 1989 yılında Bernard Baars tarafından öne sürülmüş olan “kapsamlı çalışma alanı modeli”dir. Bu model atmış yıllık psikolojik modellemenin yoğun bir sentezidir. (Deheane, 2018, s: 211) Baars, belli bir bilgi parçasının farkında oluşumuzun, dolayısıyla belli bir bilgi parçasının bilinçli olarak algılanmasının, söz konusu bilgi parçasının diğer bilişsel mekanizmaları tarafından kullanılabilir hale getirdiğini savunmuştur. Bilinçli zihnin sınırlı kapasitesine dikkat çeken Baars, beynin bilinçsiz işlem kapasitesinin

genişliğiyle kıyaslandığında bilinçli kapasitenin ilk bakışta organizma için bir dezavantaj gibi gözüktüğünü ancak bilincin evrimde bir işleve hizmet etmek amacıyla ortaya çıkmış olması gerektiğini belirtir. Bilincin sağladığı avantaj beynin birçok modülünün yaptığı hesaplamaları birleştirmek ve üst düzey hesaplamalar için söz konusu bilgiyi kullanılır şekilde temsil etmektir. Kapsamlı çalışma alanı modeli doğruysa, bilinç modülerliği azaltmak üzere evrimleşmiş olmalıdır. Bilgi, bütünsel nöron çalışma alanı sayesinde beynimizin modüler işlemcileri arasında serbest bir şekilde paylaşılabilir. Bilginin bütünsel olarak kullanılabilirliği, tam da öznel olarak yaşadığımız bilinç halidir. (Deheane, 2018, s: 218) Özellikle tiyatro metaforu ile bilinen “kapsamlı çalışma alanı modeli”ni diğer modellerden ayıran en önemli özelliği, bilince, beynin alt ve üst seviye bilişsel aktivitelerinin arasında işlevsel bir özellik affediyor olmasıdır. Baars bilinci betimlemek için bir tiyatro metaforu kullanmıştır. Bu metaforda bilinç, seçici dikkat olarak betimlenen sahne ışığının aydınlattığı tiyatro oyuncusudur. Oyuncu diğer oyuncularla etkileşimlere girmekte ve rolünü oynayarak bilincin içeriklerini ortaya çıkarmaktadır. Ancak sahnenin gerisinde oyunu yazan senarist ve oyunun sergilenmesini sağlayan birçok çalışan vardır ki bu çalışanlar bilinçli deneyimi ortaya çıkaran değişik nöronal aktivite gruplarını temsil ederler. Oyunu seyreden seyirciler de karanlıkta kalmıştır, dolayısıyla bilinçsiz üst seviye bilişsel işlevleri temsil ederler. (Baars, 1997, s: 40-42) Deheane, Baars’ın global çalışma alanı modelini temel alarak bu modeli bilinçli deneyimin nöronal korelasyonlarını araştırarak sınamış ve geliştirmiştir. Temelde Baars’ın kapsamlı çalışma alanı modeli ampirik gözlemler sonucu şekillenmiş, üç önerme üstüne inşa edilmiştir. İlk önerme beyindeki bilişsel süreçlerin çoğunun bilince gerek duymaksızın gerçekleşebileceğini öne sürer. Bilinçli deneyim oluşturacak eşiğin altında kalan duyusal bilgiler beyin tarafından işlenebilir ve bilişsel sistemlerde farklılıklar yaratabilirler. İkinci önerme seçici dikkatin bilinçli deneyim için ön koşul olduğudur. (Dikkatin iki farklı türü vardır. Ani bir sese yöneltilen dikkat beyinde hiyerarşik olarak aşağıdan yukarı bir süreç, dikkatin istemli bir şekilde belli bir göreve yönlendirilmesi olarak tanımlanan seçici dikkat ise yukarıdan aşağı bir süreçle tanımlanırlar.) Üçüncü önerme ise konuşma, istemli hareket, karar verme, uzun süreli bellek gibi üst düzey bilişsel işlevler için bilincin gerekli olduğudur. (Dehenae, Naccache, 2001, s: 5-8)

Kapsamlı çalışma alanı modeline göre, herhangi bir zamanda korteksteki modüler yapılar birbirileriyle paralel olarak bilgiyi bilinçli deneyim yaratmaksızın işlerler. Alt seviye bilişsel işlevleri yerine getiren modüllerde herhangi bir bilinçli deneyim oluşmaz. Örneğin V1 bölgesindeki seçici nöronlar her zaman için bir çizgiyle

karşılaştıklarında yoğun şekilde ateşlenirler. İşlenmekte olan belirli bir bilginin bilinçli deneyim yaratması ancak birbiriyle paralel işleyen modüllerdeki bilginin seçici dikkat mekanizmaları tarafından amplifiye edilerek tüm kortekse yayılmasıyla mümkündür. Seçici dikkat bir çizgiye yönlenmişse V1 nöronlarının bilgisi diğer modüller ile paylaşılarak çizginin görüntüsü bilinçli bir görsel deneyime dönüşür. Kapsamlı nöronal çalışma alanı modelinde belli bir bilginin tüm kortekse yayılması anatomik olarak belirlenmiş nöronlar; çalışma alanı nöronları aracılığıyla gerçekleşir. Görüntüleme teknikleri de beynin değişik modüllerinin arasında bilgi taşıyan uzun akson uzantılı nöronların varlığını doğrulamaktadır. (Dehenae, Naccache, 2001, s: 26) Söz konusu bilgi, diğer modüller ile paylaşıldığında hem bilinçli deneyim yaratır, hem de üst modüller tarafından kullanılabilir hale gelir. Dolayısıyla kapsamlı çalışma alanı modeline göre bilgi global olarak kullanılabilir şekilde korteksin değişik modüllerine dağıtıldığında, bu bilgi aynı zamanda öznel bilinçli deneyimi yaratmaktadır. Yine aynı şekilde öznel bilinçli deneyim yaratan bilgi, uzun süreli bellek oluşturmak, öznel deneyimlerin rapor edilmesi, planlama ve istemli hareketler gibi üst düzey bilişsel mekanizmalar tarafından kullanılabilir hale gelmektedir. Deheane, kör görüşünden mustarip hastaların koridorda objelere çarpmadan yürüyebildiğini, hatta koridordaki obje sayısını tahmin etmeleri istendiğinde hastaların cevaplarının tahmin edilemeyecek kadar isabetli olduğunu, dolayısıyla bu hastalarda retinadaki bilginin V1 üstünden olmasa da beynin belli bölgelerine ulaşarak belli ölçüde işlenmekte olduğunu öne sürer. İşlenen bilgi beynin birçok alt modül tarafından kullanılmasına rağmen kör görüşü hastaları hiçbir zaman objelerden birinin yerini değiştirmek gibi istemli bir harekette bulunmazlar. Deheane, birçok bilişsel sistemin işlevi sonucu çıkan istemli hareketler içinde bilinçli algının ön koşul olduğunu belirtir. (Deheane, Naccache, 2001, s: 11)

Deheane, alt modüllerin işlediği ancak bilince çıkmayan bilgilerin beyin aktivitelerinde nasıl farklılıklar yarattığını anlamak ve bilinçli deneyimin nöronal korelasyonlarını incelemek için maskeleme tekniklerine dayanan deneyler tasarlamıştır. Bir görüntünün bilinçli algılanabilme süresi ortalama 500 ms'dir. Maskeleme deneylerinde deneğe sırasıyla 500 ms gürültüden oluşan bir ekran, 50 ms bir kişinin fotoğrafı, 50ms gürültülü ekran ve 500 ms bir kişinin fotoğrafı gösterilir. Gösterilen fotoğraflardan ilki bilince ulaşamaz, ikinci fotoğrafın gösterilme süresi ise görsel bilginin beyinde bilinçli deneyim oluşturulması için yeterlidir. Eğer bilinç seviyesine ulaşmayan fotoğraf bilinç seviyesine ulaşan fotoğraf ile aynıysa deneklerin fotoğrafı tanıma süreleri, bilinç seviyesine ulaşmayan fotoğrafın bilinçli deneyimlenen fotoğraftan farklı olduğu

durumdan daha kısa zaman alır. Deheane, bu deneyde beyindeki yüz tanıma modüllerinin gelen çelişkiyi ayırt edebilmek için daha fazla aktiviteye dahil olduğunu ve fazla aktivitenin deneyin tepki süresini uzattığını belirtir. (Deheane, 2018, s: 57-60) Bilinçli deneyim yaratmamasına rağmen kısa bir süre boyunca gösterilen fotoğrafların beyinde yanıtlama süresini etkileyen bir aktiviteye neden oluyor olması, bilince erişimi olmayan görsel bilginin beyinde işlendiğinin kanıtıdır. Aynı maskeleme deneyinde deneğe bilinç eşiğinin altında öfkeli bir insan fotoğrafı verildiğinde, deneğin amigdalasında hatırı sayılır bir aktivite gözlenmiş, ancak denek öfkeli insan fotoğrafını bilinçli olarak görmediğini belirtmiştir. Dolayısıyla duyusal girdiler bilinçli deneyim yaratmaksızın beynin değişik modüllerinde işlenebilmektedir. Belli bir duyusal bilginin bilinçli deneyim yaratması için belli bir süre boyunca beyindeki belli modülleri aktive etmesi gerekmektedir. Ancak belli bir yoğunlukta ve yeterli bir süre boyunca aktif olan duyusal bir girdi de illaki bilinçli bir deneyim oluşturmaz.

Dikkat bilincin ön koşuludur, ancak dikkat ve bilinç aynı mekanizmalar tarafından tanımlanmaz. Koch, dikkat mekanizmalarının bilinçle ortak birçok nöron devresi kullanıyor olmasının olası olduğunu ancak dikkat ve bilincin nöronal mekanizmalarının ayırt edilmesinin önemli olduğunu vurgulamıştır. (Koch, 2004, s: 154) Primatlarda retinadan LGN'ye bilgi taşıyan ortalama bir milyon ganglion hücresi vardır, bu ganglion hücreleri saniyede bir mega bitlik bilgi taşıma kapasitesine sahiptirler. Beynin gerçek zamanda işleyebileceği bilgi miktarıysa saniyede bir mega bittin çok altındadır. Ayrıca beyin görsel uyaranların bilgisi dışında diğer duyusal uyaranların ve vücudun içsel durumlarının bilgisine de erişmelidir. Dikkatin evrimde önemli olan bilginin seçilip, önemli olmayan bilginin işlenmesini engellemek için geliştiği düşünülmektedir. (Koch, 2004, s: 153) Tüm omurgalıların beyni dikkat mekanizmalarına sahiptir. Beyinde iki farklı dikkat mekanizması bulunmaktadır. Aşağıdan yukarı dikkat, çevre tarafından belirlenir. Görsel sistem kesik hareketlerle çevreyi otomatik olarak tarar. Görsel alan içindeki düzensizlikler ya da ani değişimler bu tarama esnasında aşağıdan yukarı dikkat mekanizmalarını harekete geçirir. Aşağıdan yukarıya dikkat ani bir uyarana yöneltilen dikkattir. Seçici dikkat esnasında ise beyinde yukarıdan aşağı bilgi akışı yoğundur, seçici dikkat mekanizmaları aktif iken beynin üst düzeyde bilgi işleyen bölgelerinden alt modüllere yönelen bilgi akışı, alt modüllerin hangi bilgiyi işlemesi gerektiğini tayin eder.

Dikkatin bilincin ön koşulu olduğunu destekleyen en ilginç deneylerden biri "görünmez goril deneyi"dir. Bu deneyde, deneklerden iki basket takımının aralarında kaç pas

yaptıklarını saymaları istenir. Deney devam ederken belli bir süre boyunca goril kostümü giymiş bir oyuncu paslaşan oyuncular arasında dolaşır, elini kolunu sallar, sonra sahneyi terk eder. Deneyin sonunda deneklerin hatırı sayılır bir yüzdesi, goril kostümü giymiş olan oyuncuyu görmediklerini beyan etmişlerdir. (Deheane, 2018, s: 54) Seçici dikkat mekanizmaları yapılması gereken göreve göre, görevin dışında kalan tüm uyaranları inhibe etme özelliğine sahiptir. Dolayısıyla duyusal uyaranların bilinçli deneyim yaratması ancak dikkat mekanizmalarının o uyaranları seçmesi ile mümkün olmaktadır. Yapılmış olan görüntüleme çalışmaları dikkat edilmeyen duyusal girdilerin beyinde hatırı sayılır bir aktivite yaratmadan bir süre boyunca sadece belli modülleri aktive edebildiğini ve kısa sürede bu aktivitenin yok olduğunu göstermiştir. Dolayısıyla ancak dikkat edilen uyaranlar bilinçli deneyim yaratırlar. Bilinçli deneyim, söz konusu bilginin çalışma alanı nöronlarınca beynin birçok modülüne ulaşmasıyla oluşur. Bilince ulaşabilen bilgiler ise üst düzey modüller için kullanılabilir hale gelirler. Evrim esnasında av ve avcı ilişkisinin baş göstermesiyle avcının avını yakalamak, avın ise avcıdan kaçmak için planlama yetisi elzem hale gelmiştir. (Mclver, 2017, s: 2381) Dolayısıyla çevredeki önemli bilgilerin seçilmesi av veya avcının planlama yapmaları için gerekli olan gürültüden ayıklanmış bilgiye ulaşmaları için şarttır. Kapsamlı çalışma alanı modeline göre, bilinç planlama yapmak için gerekli duyusal girdilerin özetini çıkarmak için evrimleşmiştir. Eğer çevrede çok fazla gürültülü bilgi varsa zombi sistemler gerekli tepkilerin verilmesi için yeterli olamazlar. Öğrendiğimiz birçok aktivite belli bir süre sonra zombi sistemler tarafından yerine getirilmeye başlanır ancak yeni bir durumla karşı karşıya olduğumuzda, duruma uygun cevabı verecek bir zombi sistemden yoksunsak, beynin en uygun cevabı üretmesi ancak birçok modülün etkileşimiyle mümkün olur. Dolayısıyla, bilinç ve bilincin bilgi paylaşma özelliği, beynin en uygun cevabı üretmesi için gereklidir.

“Stroop test” olarak tanımlanan deneyde, fontları ters renkle yazılmış olan sarı ve yeşil kelimelerini seçmeleri istenen denekler, sarı ve yeşil kelimeleri yeşil ve sarı fontlarıyla yazılmışken bu kelimeleri seçerken kelimeler ve renklerin uyumlu olduğu versiyona göre oldukça yavaş davranmıştırlar. Bunun nedeni, zombi sistemlerin birbiriyle çelişkiye düşmesi sonucunda dikkat mekanizmalarının kelime tanıma modüllerini aktive ederken, renk bilgisini işleyen görsel sistemin, özellikle V4’ün aktivitesini inhibe etmesi, dolayısıyla beynin çelişkiyi giderebilmek için daha az modül ile işlem yapmasıdır. (Deheane, Naccache, 2001 s: 10-11)

Deheane, bilincin tam bir teorisinin beyin aktivitelerinden neden bazılarının geçici veya kalıcı şekilde bilinçli erişimi olmadığını, bilinç içeriklerinin tümünün neler olabileceğini ve bu bilinçli içeriklere beyindeki hangi nöron aktivitelerinin neden olduğunu açıklaması gerektiğini vurgular. (Deheane, Naccahe, 2001, s: 15) Deheane'e göre, beynin ancak çalışma alanı nöronlarıyla birbirine bağlanmış modülleri bilinçli deneyim oluşturur. Çalışma alanı nöronları çift taraflı bilgi iletirler ve değişik modüller arasında iletişimi sağlarlar. Bu nöronları içeren modüller duyusal algıları işleyen, istemli motor hareketlerini kontrol eden ve uzun süreli bellek için gerekli olan modüllerdir. Deheane ayrıca prefrontal korteks ve korpus collusumun çevresinde, frontal lobun alt kısmında bulunan anterior cingular korteksinde çok fazla sayıda çalışma alanı nöronu içerdiğini ve bu bölgelerin insanlarda karar verme ve kendini değerlendirme gibi işlevleriyle öne çıktığının altını çizer. (Deheane, Naccahe, 2001, s: 24, 26) Bu iki bölgenin çok çeşitli ve farklı bilinçli deneyimlerde aktive olduğu gözlemlenmiştir. Anterior cingular korteks özellikle hislerin zihinsel temsiller ile ilişkisini belirler. Ancak bilinçli deneyimin beyinde bir yeri yoktur. Bilincin içerikleri birbirinden çok farklı ve sürekli değişen içeriklerdir. Dolayısıyla bilinçli bir deneyimde, dikkat mekanizmalarının kontrolünde belli modüller aktive olurken belli modüller kapsamlı çalışma alanından çıkarlar. Kapsamlı çalışma alanı her bilinçli deneyimde ve bilinçli bir deneyimin içinde değişen modüllerin aktivasyonuyla devamlı değişmektedir. Dinamik seferberlik olarak tanımlanan bu aktivitede bilinçli deneyimi tarif eden bir aktivasyon deseni yoktur. Değişik bilinçli deneyimler değişik modülerin seferberliğini gerektirmekte ve her bilinçli deneyim kortekste çok farklı aktivasyon desenleri oluşturmaktadır. (Deheane, Naccahe, 2001, s: 16)

### 2.5.2 Entegre Bilgi Teorisi

2004 yılında Guiliano Tonini, nöronal mekanizmalardan ya da ampirik gözlemlerden bilinçli deneyimin yapısına ulaşmaya çalışan nörobilim modellerinden farklı bir şekilde bilinçli deneyimlerin özellikleriyle başlayarak bilincin nöronal mekanizmalarına ulaşmayı hedefleyen entegre bilgi teorisini yayınlamıştır. Entegre bilgi teorisi, nörobilimin modellerinden farklı bir yaklaşımla bilinci açıklamaya çalışıyor olsa da entegre bilgi teorisinde, bilinçli içeriklerin bilgilendirici olduğu ve bilginin üretiminin sistem içinde fark yaratan farklılaşmalarla mümkün olduğu belirtilmiştir.

Entegre bilgi teorisi ilk olarak birbirinden bağımsız beş aksiyomla olası bir bilinçli deneyimin sahip olduğu tüm fenomenolojik özellikleri tanımlar. Daha sonra bilinçli deneyimi tanımlayan aksiyomlardan bilinçli deneyimi yaratabilecek fiziksel mekanizmanın özelliklerini türetir. Bilinci fenomenolojik olarak tanımlayan beş aksiyom varoluş, kompozisyon, bilgi, bütünleşme ve dışarıda bırakmadır.

İlk aksiyom, varoluş aksiyomu, ancak var olanın bir fark yaratabileceğini, var olmayanın ise fark yaratamayacağını belirtir. Fiziksel bir şeyin var olabilmesi için çevresinde ya da kendi içinde fark yaratabilmesi lazımdır. Ancak fark yaratan bir şey var olabilir, bir şeyin deneysel olarak gözlemlenebilmesi dahi o şeyin deneysel düzenekle etkileşime girmesiyle, dolayısıyla deneysel düzenekte bir fark yaratabilmesiyle mümkündür. Bilinçli deneyim birinci sahsın deneyimidir ve üçüncü sahsın gözlemiyle direkt olarak ulaşılamazdır. Bilinç üçüncü şahsın gözlemine kapalı olduğundan, bilincin içeriğini oluşturacak farklılaşma birinci sahsın kendi sistemi içinde, birinci şahıs açısından fark yaratabilmelidir. İkinci aksiyom, kompozisyon, bilinçli deneyimin birçok bileşenden oluşan bütünlüklü bir yapısı olduğunu belirtir. Bilinçli bir deneyimi oluşturan birçok alt küme bulunur, bir cismin şekli ve rengi o cismin bilinçli algısının iki alt kümesidir ve bilinçli algının sadece şekil ya da renkten oluşması mümkün değildir. Bilinçli her deneyim birbirinden ayrılamayan çok çeşitli alt kümelerden oluşur. Üçüncü aksiyom olan bilgi aksiyomu, bilinçli deneyimin doğası gereği çok çeşitli bilgi içerikleri olduğunu, dolayısıyla çok farklı ve çeşitli bilgi içeriklerinin her bilinçli deneyimi bir başka bilinçli deneyimden ayırt edeceğini belirtir. Dördüncü aksiyon olan bütünleşme, bilinçli deneyimin bütünlüğünü tanımlar. Bilinçli deneyim çeşitli duyusal algılar, düşünceler ve duygulardan oluşur ve bilinçli deneyim onu oluşturan her parçayı bütün bir şekilde birinci şahsa sunar. Beşinci aksiyom olan dışarıda bırakma ise bilinçli deneyimin bütünlüklü yapısının sınırları olduğunu belirtir. Bilinçli deneyimin içerikleri çok çeşitlidir ancak bu içerikler belli sınırlara sahiptir. Örneğin bir insanın bilinçli deneyiminde kendi kan basıncına erişiminin olması mümkün değildir. Bir bilinçli deneyim aynı zamanda belirli bir zaman ve belirli bir mekanda gerçekleşir. Bilinçli deneyim, farklı bilinçli deneyimlerin birleşimi olarak tanımlanamaz. (Tonini, 2015, s: 5-6)

Tonini'nin öne sürmüş olduğu beş aksiyom, bilincin tanımlayan birbirinden bağımsız beş özelliğidir. Bilincin kendisini aksiyomlar ile tanımlayan entegre bilgi teorisi, bilincin ancak fiziksel bir mekanizmanın ürünü olduğunu ve söz konusu mekanizmanın fiziksel kurallarının bilincin beş aksiyomuyla uyumlu olması gerektiğini öne sürer.

Dolayısıyla beş aksiyoma denk gelen beş postulat bilincin fiziksel temelindeki mekanizmaların sahip olması gereken beş özelliktir. (Tonini, 2015, s: 6)

İlk postulat varoluş postulatıdır ve bilinçli deneyimin var olabilmesi için bilinçli deneyimi deneyimleyen maddesel bir mekanizma olması gerektiğini belirtir. Aynı zamanda bilincin var olmak için fark yaratmasının gerekliliğine tekabül eden postulatta, bilincin fiziksel sisteminin kendi içinde fark yaratabilmesini sağlayacak etki tepki gücü olması gerektiği belirtilir. Beynin söz konusu bilinçli deneyimi yarattığı düşünülürse, bu postulat, beyinde bilinçli deneyime denk gelen bir durum olduğunu ve bu durumun geçmiş durumlara ve gelecekte oluşma olasılığı olan durumlara etki tepki yapılarıyla bağlı olduğunu açıklamaktadır. Bilinç dış gözlemcinin ulaşımına kapalı olduğundan etki  tepki yapıları beynin kendi içindeki yapılardır. Bilinçli deneyim, beynin içindeki etki tepki mekanizmalarının bir sonucudur. İkinci postulat bilincin kompozisyon aksiyomuna tekabül eden postulattır. Bilinçli deneyimin bütününü oluşturan sistem, bilinçli deneyimin içeriklerini oluşturan alt sistemlerden oluşur. Bütün sistemin bir yapısı vardır ve alt sistemlerin işleyişleri ve nedensellikleri ana sistemin yapısında belirlenmiştir. Üçüncü postulat ise bilgi aksiyomu ile alakalıdır ve bilinçli bir deneyimin, olası birçok deneyimden ayırt edilebilir olduğunu belirtir. Bir andaki bilinçli deneyimin içeriği, o deneyimin geçmişindeki birçok olası duruma ve deneyimin geleceğindeki birçok olası duruma etki tepki yapısıyla bağlıdır, ancak birçok olası durumdan sadece bir tanesiyle tanımlanır. Deneyim ne kadar çok bilgi taşıyorsa, deneyimi oluşturan fiziksel sistemin entropisi o denli yüksektir. Dördüncü postulat bilinçli deneyimin bütünlüğü aksiyomuna tekabül eder. Bütünlük postulatına göre bilinçli deneyimi yaratan fiziksel durumun başka fiziksel durumların toplamına indirgenemez olması gerekir. Bütün bir bilinçli deneyim, onu yaratan birbirinden bağımsız deneyimlerin toplamı değildir, bilinçli deneyimi yaratan her içerik birbirilerine etki tepki yapıları ile bağlıdır ve bir fiziksel sistemin bölünemez işleyişi içinde tanımlanmalıdırlar. Eğer bir sistem iki sistemin toplamı olarak tanımlanabiliyorsa, söz konusu sistemin bütünlüğünden bahsedilemez. Ancak sistemde iki bağımsız sistem oluşmuşsa, bilinçli deneyimi yaratan parça, daha fazla bilgi taşıyan, entropisi yüksek olan parça olacaktır. Beşinci postulat ise dışta bırakma postulatıdır ve bilinçli deneyimle ilişkili söz konusu fiziksel sistemin alt sistemlere indirgenemez en zengin içeriğe sahip parçasının uzay zamanda tek bir etki tepki yapısı ile tanımlanabilir olduğunu belirtir.

Entegre bilgi teorisinin beş postulatından kendi içinde etki tepki yapısına sahip, dolayısıyla kendinde fark yaratabilen, başka sistemlere indirgenemez bir sistemin bilinçli deneyimler yaratabileceği sonucu çıkmaktadır. Başka bir ifadeyle, belirli bir entropiye sahip, dolayısıyla bilgi taşıyan ve yapısında etki tepki mekanizmalarıyla zamanda kendi içinde değişiklik meydana getirebilen, kendinden basit sistemlere bölünemeyen her sistemin kendi bilinçli deneyimi olma olasılığı vardır. Entegre bilgi teorisi bilinçli deneyimi deneyimleyen sistemin maddesinin türünden bağımsız bir teoridir, canlı veya cansız sistemlere uygulanabilir. Bir sistemin içinde bulunduğu tek bir durum, geçmişindeki ve geleceğindeki durumlar cinsinden yüksek boyutlu uzayın bir boyut olarak temsil edilir. Entegre bilgi teorisi geometrik bir teoridir ve her durum yüksek boyutlu uzayın bir boyutunu oluşturur. Söz konusu indirgenemez bir durumun taşıdığı bilgi, etki tepki yapısına bağlı olarak geçmiş ve gelecek uzayının oluşturduğu durumlarının toplam boyutu olarak hesaplanır. Bu hesabın sonucunda ulaşılan phi sayısı, sistemin taşıdığı entegre edilmiş bilgiyle orantılıdır. Phi sayısı bir sistemdeki en büyük indirgenemez parçanın durum uzayının boyutun ölçüsüdür. Bilinçli bir sistemde phi her zaman sıfırdan büyüktür. (Tonini, 2015, s: 9) Örneğin bir durumu oluşturabilecek sekiz durum varsa ve andaki durum sekiz değişik gelecek durum yaratabiliyorsa, bu sistem on altı boyutlu bir sistemdir. Etki tepki yapısı geçmiş durumlar ile andaki durum ve andaki durumun oluşturabileceği gelecek durumlar ile andaki durum arasındaki ilişkiyi veren fonksiyonlar olarak ifade edilir. Sistemin kendi içinde fark yaratması geçmiş ve gelecek durumların birbirini etkilemesine bağlıdır.

İleri beslemeli bir sistemde bu fark her zaman sıfır olacak iken döngüsel, geri bildirimli bir sistemde geçmiş ve gelecek durumlar birbirleriyle ilişki içindedir. Eğer geri bildirimli bir sistem alt sistemlere indirgenemez ise Phi sıfırdan büyük olacaktır.

Entegre bilgi teorisi, neden korteksten daha fazla sayıda nöron içeren beyinciğin bilinçli deneyim yaratmadığını, beyincikteki nöronal bağlantıların çoğunlukla ileri beslemeli bir sistem oluşturduğu için beyinciğin etki tepki yapısına sahip bir sistem olmadığını, dolayısıyla andaki durumun, geçmiş ve gelecekteki olası durumlarla bağlantısının bütünleşik çok boyutlu bir uzay yaratmasının söz konusu olmadığını belirterek açıklar. Beyincik adım adım çok karmaşık hesaplar yapar ama geri bildirim yapıları olmadığından kendi içeriğinde kendisi fark yaratamaz. Entegre bilgi teorisinden çıkan başka bir sonuç ise ileri beslemeli yapay sinir ağlarının, örneğin görüntü tanıyan evrişimsel sinir ağlarının bilinçli deneyim yaratmasının mümkün

olmadığıdır. Ancak döngüsel yapıya sahip olan ve kendi içeriklerindeki bilgide fark yaratan tekrarlayan sinir ağlarının Phi değeri sıfırdan büyüktür.

Phi sayısı bilinç seviyesinin bir ölçütüdür. Bilinçli deneyimin kalitesi, bilinçli deneyimin içeriği ise bilinçli deneyimi yaratan mekanizmanın birbiriyle etkileşebilen geçmişteki ve gelecekteki tüm durumların çok boyutlu durum uzayında oluşturduğu şekle tekabül eder. Örneğin beyin gibi nöron devrelerinden oluşan bir yapıda, bilinç içeriği o andaki nöronal aktivasyonu yaratan ve o andaki norönal aktivasyonun yaratabileceği aktivasyonların yarattığı uzayda birbiriyle sıfırdan farklı bir fonksiyonla ilişkilendirilebilen tüm aktivasyonların şekliyle belirlenecektir. Bu şekil ise bilinçli deneyimin nöronal korelasyonlarının oluşturduğu aktivasyon desenine denk gelecektir. (Tonini, 2015, s: 12) Dolayısıyla entegre bilgi teorisi insan beynine uygulanılabilirse beyinde değişik nöronal aktivasyon şekillerinden kişinin kırmızı mı yoksa yeşil mi gördüğü anlaşılabilecektir.

Entegre bilgi teorisi pratikte çok basit, belli sayıda olası durumu olan sistemlerde dahi çok fazla ve karışık hesaplamalar yapılmasını gerektiren bir teoridir. Ancak birçok varsayımla basitleştirilmiş bir versiyonu koma durumundaki hastaların durumlarının değerlendirilmesi için kullanılmakta ve hastaların bilinç seviyesine dair tutarlı sonuçlar vermektedir. (Tonini, 2015, s: 15)

## 2.6 HİSSEDEN ZİHİN VE PEKİŞTİRMELİ ÖGRENME

Nörobilim modelleri bilinçli deneyimleri oluşturan mekanizmaları araştırırken bilinçli deneyimle doğrudan alakalı yeni kortekse, yeni kortekste bulunan karmaşık nöron devrelere yoğunlaşmıştır. Yeni korteksin bilinçli deneyimle inkar edilemez bir korelasyonu olduğu açıktır. Ancak Damasio gibi bazı bilim insanları bilincin sadece nöron devreleri ya da bilgi teorileri ile açıklanamayacağını savunur. Damasio'nun temel noktası duygulardan evrimleşmiş hislerin bilinçli deneyimle sıkı bir etkileşim içinde olduğudur. Hisleri bilinç olmaksızın, bilinci ise hisler olmaksızın anlamak mümkün değildir.

Beyin sapının üst kısımlarında, aralarında hipotalamusun da bulunduğu birçok çekirdek yapı bulunmaktadır. Bu çekirdeklerden en önemlisi olan lokus seruleus bin kadar nörondan oluşmaktadır. Beyin sapının arka tarafında bulunan lokus seruleusun zarar görmesi geri dönüşsüz bilinç kaybına neden olur. Beyin sapı çevresindeki tüm

çekirdeklerin nöronları dağınık bir şekilde beynin tamamına dağılmışlardır. Bu çekirdeklerdeki nöronlar, ağrı, ani ses, ani ışık gibi beklenmedik durumlarda ve vücudundaki beklenmedik değişikliklerde ateşlenerek beyinde nöro-modülatörlerin salgılanmasına neden olurlar. Nöro-modülatörler sinapsların hem gücünü, hem de bağlantılarını değiştirme gücüne sahip kimyasallardır. Bu nedenle beyin sapı çevresindeki çekirdek sistemlerine değer sistemleri denilmektedir. (Edelman, Tononi, 2019, s: 71)

Damasio, bilincin duygulardan bağımsız araştırılmaması gerektiğini savunur. Beyinler bilinçli zihinleri korteks düzeyindeki değil, beyin sapı düzeyindeki değer sistemleri aracılığıyla oluşturmaya başlar. Çoğu duygusal tepki beyin sapında şekillenir. (Damasio, 2020, s: 23) Crick ve Koch, beyin sapı, *hipotalamus,* bazal *ganglia*, kısacası bu bölgelere dağılmış olan ve değer sistemlerini veya değer sistemleriyle paralel çalışan yapıları bilinçli deneyimler için arka plan koşulları olarak tanımlamıştır. Ancak Damasio evrimsel olarak ilkel olan bu bölgenin bilincin temeli olduğunu ve kortekste gerçekleşen her türlü nöronal işleme anlam katacak değerlerin bu bölgeler tarafından üretildiğini vurgular.

### 2.6.1 Damasio'nun Benliğe Dayalı Bilinç Modeli

Bilinç problemine evrimsel bir açıdan yaklaşan Damasio, beyin korteksinin hatta sinir sisteminin ortaya çıkmasından önce basit organizmaların hayatta kalmasını sağlayan mekanizmalar geliştirmiş olduğunu belirtir. (Damasio, 2020, s: 27) Bu mekanizmalar organizmanın kötüden sakınmasını ve iyi olana yönelmesini sağlayan duygu mekanizmalarıdır. Duygu mekanizmaları, DNA yapısının dizilimi sayesinde yönettikleri kimyasal tepkimelerle tek hücreli bakterilerde dahi çok çeşitli davranış repertuarları, hatta sosyal davranış repertuarları oluştururlar. Bir bakterideki hayatta kalma mekanizmaları tamamen genler tarafından tanımlanmıştır. Genler yaşamın şifresidir ve dengeleşim yeteneğinin ortaya çıkmasından sorumludurlar. (Damasio, 2020, s: 36-37)

Damasio duygu ve his arasında bir sınır çizer, duygular otomatik mekanizmalardır ve her duygu organizmanın hareketi veya değişik kimyasallar salgılaması gibi bir aktiviteyle bağlantılıdır. His ise sinir sistemine sahip canlılara özgüdür, ancak hisler duygulardan evrimleşmiştir. Damasio, canlıların zihinleri olmadan önce de zihni ve bilinci olan canlıları anımsatan davranışlar sergilemiş olduklarının altını çizer. Bir

bakteri dahi çevresini algılayıp, tepki verebileceği donanımlara sahiptir. Tek hücreli canlılardan çok hücreli canlılara geçilip hücre koalisyonları kurulup, bu koalisyonlar karmaşıklaşarak bir sinir sistemi oluşturduğunda, bakterideki etki tepki mekanizmalarıyla çalışan duygular, bilinçle paralel olarak beyindeki haritalarda duyguların hisleri olarak temsil edilmiş, bu temsiller dış dünyanın temsilleriyle birleştirilmiştir. (Damasio, 2018, s: 94) Bir duygunun hisse dönüşmesi, o hissin hissedilebilmesini gerektirmektedir, dolayısıyla his ve bilinç evrimde birbirine paralel gelişmiştir. Nasıl ki ilkel organizmalarda etki tepki mekanizmalarının hedefi dengeleşimi korumaksa, en karmaşık organizmalar da temelde dengeleşimlerini korumak amacını taşırlar. Gelişmiş bir zihin, sadece anlık uyaranların dışına çıkarak geleceği öngörecek sistemleri dengeleşimini korumak için kullanır. Damasio'ya göre insani yaşama arzumuz, bedenimizdeki biyolojik değer taşıyan tüm hücrelerin ham iradelerinin bir toplamıyla ortaya çıkmış olabilir. (Damasio, 2020, s: 37)

Damasio bilinçli bir zihnin temelinde bedenin olduğunu vurgular. Beyin içinde bulunduğu vücuda hizmet eder ve beynin esas fonksiyonu karmaşık organizmaların dengeleşimini sağlamaktır. Beyne devamlı olarak vücudun içsel durumları ile ilgili bilgi iletilmektedir. Beyinde aynen duyusal girdiler ile şekillenmiş temsiller olduğu gibi vücudun içsel durumlarının nöronal temsilleri de bulunmaktadır. Vücudun içsel mekanizmalarında oluşan durumlar, acıkmak, susamak, halsiz hissetmek gibi *qualia* boyutunda hislere denk gelir. Organizmanın hislere kayıtsız kalma şansı yoktur, eğer ağrınız varsa o ağrıyı görmezden gelemezsiniz. Acı ve haz arasındaki skalada duygular organizmaya ne yapması gerektiğini bildirir. Vücudun iç temsili olan duygular, o duyguyu oluşturan dış temsiller ile bağlantı kurduğunda hisler meydana çıkar. His, belirli bir bedensel durum algısının, belirli konularla ilgili düşüncelerle birlikte algılanmasıdır. (Damasio, 2020, s: 90)

Damasio, beyin hasarları nedeniyle değişik hisleri kaybeden hastalar üstüne yaptığı araştırmada, değişik hislerin beynin değişik bölgelerinde temsil edildiğini, bir hasar sonucu bir hissin kaybı yaşanırken diğer hislerin korunabildiğini belirtir. Hisler ile ilgili en şaşırtıcı araştırmalar, hisler ile üst düzey bilişsel mekanizmalar arasındaki ilişkiler üstüne yapılmış çalışmalardır. Örneğin hastalar belli bir duyguyu ifade etme yeteneğini kaybettiklerinde, aynı zamanda onu uyandıran hissin denetimi de kaybederler. (Damasio, 2018, s: 6) Duygular ve bilişsel yapılar birbirileriyle ayrılmaz şekilde bir bütünün parçalarıdır. İnsanlarda, hislerin tamamen rasyonel bir mekanizma olduğu düşünülen karar verme mekanizmasıyla ilişkisini inceleyen Damasio, önceden

normal olan bireylerin belli beyin bölgeleri hasar gördüğünde bu hastaların toplumsal ve finansal hayatlarını yönlendirme becerilerini kaybettiğini gözlemlemiştir**.** (Damasio, 2018, s: 144) Özellikle sol prefrontal lobta hasarı olan hastalar, bilişsel kapasitelerinde hiçbir düşüş yaşamaz ama bu hastalar basit kararlar vermekte bile zorlanırlar. Bunun nedeni bu hastaların eylemleri duygularla ilişkilendirme yeteneklerini kaybetmiş olmalarıdır. Prefrontal korteks belirli bir eylem ve buna bağlı duygusal sonucun öğrenilmesinden, dolayısıyla eylemi ödül veya ceza duygularıyla bağdaştırmaktan sorumludur. Bu bölgedeki hasar, eylemlerin ve eylemlerin olası sonuçlarının duygusal temsillerinin bağlantısını kopardığından, hasta olası bir eyleminin sonuçlarını tartamaz, eylemde bulunması veya bulunmaması için ona yön gösterecek duygu ve bu duygulardan türetilmiş hislerden yoksundur.

Damasio bilinci benlik çerçevesinde araştırmıştır. Damasio, benliği, bağlantılı zihinsel süreçlerin toplamında ifade bulan ve yaşayan bedenin temsiline odaklanan bağlantılı nöral süreçlerin dinamik toplamı olarak tanımlar. (Damasio, 2020, s: 10) Dolayısıyla Damasio benlik ile kendilik bilincini ayırır. Benliğin oluşumu üç aşamalı bir süreçtir. Benliğin üç aşaması ilk benlik, çekirdek benlik ve otobiyografik benliktir. (Damasio, 2020, s: 11)

Benliğin ilk aşaması olan ilk benlik değer sistemlerinde, beyin sapı ve *hipotalamus*taki çekirdek yapılarıyla şekillenen benliktir. Bu yapılar sürüngenlerde dahi bulunan yapılardır. İlk benlik organizmanın ilkel duygularını yaratır, bu duygular kabaca organizmanın vücudunda gerçekleşen süreçleri temsil eden haz ve acıdır. İlkel duygular zihinde temsil edilir ancak Damasio'ya göre bu duygulara bilincin eşlik etmesi şart değildir. (Damasio, 2020, s: 21-24)

Çekirdek benlik zihinlerin oluşumundaki ilk aşamadır. Çekirdek benlik organizma çevresindeki nesne ve olaylarla etkileşime girerken oluşur. Nesne ve olayların nöronal temsillerini kortekste yapılandırmaya başladığı esnada ilkel duygular çevreyi temsil eden nöron haritalarını ilişkilendirir. Dolayısıyla ilkel duygular olay ve nesneler ile ilişkilendirir, belirli olay ve nesnelere karşı hisler oluşmaya başlar, organizma olay ve nesnelere duygusal reaksiyonlar verir. Çekirdek benlik organizmanın yaşamı boyunca devam eden bir süreç, dış dünyanın temsilleri ile vücudun temsillerinin etkileşimidir. Beyin sapı ve çevre yapıları, *insular* korteks ve duyusal girdileri işleyen korteks yapıları çekirdek benliği kuran yapılardır ve birbirileriyle sürekli döngüsel etkileşimler içindedirler. (Damasio, 2020, s: 216-217) Çekirdek benlik, zihnin içsel ve dışsal temsillerinin organizmaya ait olmalarını sağlamaktadır.

Crick ve Koch'un bilinci Damasio'nun çekirdek benlik seviyesinde tanımladıkları iddia edilebilir. Çekirdek benlik, zihnin içsel ve dışsal temsillerini birleştirir, nesne ve olaylara anlam yükler; en basitinden çekirdek benliğe sahip organizma, belli bir nesneyi bilme hissine ve o nesnenin kendisi üstündeki etkilerini öngörme yetisine sahiptir. Nesne ve olayları, bu nesne veya olayların ilkel hisleri üstündeki etkilerini bilmek için organizma kendine ait bir *qualia* uzayı oluşturmak zorundadır. Ancak çekirdek benlikte kendilik bilinci yoktur veya çekirdek benliğe dil yeteneklerinin eşlik etmesi beklenmez. Çekirdek benlikte vücut ve korteks birbirine bağlanmıştır. Beyin sapı korteks ve vücut arasında bir köprüdür ancak beyin sapı sadece sinyalleri vücut ve korteks arasında aktaran bir yapı değildir. Değer sistemi içindeki çekirdekler, vücut ve korteks arasında döngüsel bir iletişim kurmanın yanı sıra, kendi aralarında da iletişim halindedirler. Çift taraflı sinyalleri iletmenin ötesinde modüle ederler. Damasio vücudun içsel durumlarını temsil eden *insula* ve *insula* ile yakından ilişkili *antrerior cingular* korteksin beyinde bilinçle alakalı, önemli yapılar olduğunu vurgular. *Anterior cingular* korteks his deneyimlerini etkinleştirerek *insula*ya eşlik eder. (Damasio, 2018, s: 124) Duygular açığa çıkarken söz konusu bilişsel temsiller duyguyla uyumlu şekilde çağrılır. Bu temsillerden bazıları açığa çıkan duyguya verilen bilişsel tepkilerdir.

Benliğin en üst aşaması olan otobiyografik benlik ise özellikle insanlara ve belli ölçülerde gelişmiş üst memelilere özgüdür; kendilik bilincini içerir. Kendilik bilincinin bir sonucu olarak organizma kendi otobiyografisini geçmiş deneyimlerinin hafızasıyla oluşturur ve gelecek planları yapma yetisine sahip olur. Kendilik bilinci, organizmanın korteksinde kendisinin temsilini üretebilme yeteneğidir. (Damasio, 2020, s: 218-221) Ancak insan zihninde dahi otobiyografik benliğe çekirdek benlik ve ilkel benlik eşlik etmektedir. Evrim her zaman ilkel yapılara ekleme yaparak ilerleyen bir süreçtir. Damasio, gerçek bir bilinç modeli oluşturulmasında tek hücreli canlılardan insanlara dek gerçekleşen evrim sürecindeki duygusal tepkileri yaratan kimyasal mekanizmaların göz ardı edilmemesi gerektiğini, bu mekanizmaların bilişsel yetenekleri anlamlı kılan duyguları oluşturan mekanizmalar olduğunun altını çizer. (Damasio, 2020, s: 26-27, 55-57)

### 2.6.2 Pekiştirmeli Öğrenme

Tekrarlayan yapay ağların dış dünyanın temsillerinden kendi iç temsillerini oluşturmaları ve iç temsillerini dinamik olarak güncellemeleri bu ağlarının en önemli

özelliği olmuştur. Ancak biyolojik beyinleri yapay sinir ağlarından ayıran en önemli özelliklerinden biri dış dünyadaki bilgilerin hangilerinin organizma için önemli, hangilerinin organizma için önemsiz olduğunu ayırt edebilmeleridir. Biyolojik beyinler dış dünyanın temsillerini oluştururken sadece dış dünyanın organizmayı ilgilendiren durumları ile uğraşırlar. Evrim sürecinde gelişmiş olan duygu mekanizmaları organizmaya dengeleşimini sürdürmek için dış dünyadaki hangi durumların organizmayla alakalı olduğunun ve bu durumlardan hangilerinin organizma için iyi veya kötü olduğunun bilgisini vermekte ve organizmanın bir hedefe yönelik hareket etmesini sağlamaktadır. Bu mekanizmalarını yapay sinir ağlarına uygulanması ise pekiştirmeli öğrenme alanını doğurmuştur. Pekiştirmeli öğrenme en basit tanımıyla deneme yanılma yoluyla öğrenmedir. Pekiştirmeli öğrenme bir amaç doğrultusunda çevre ile etkileşime girerek zaman içinde öğrenen ve öğrendikleri ile bir probleme yaklaşımlar geliştirebilen özerk ajanlar tasarlamayı hedeflemiştir. Günümüzde kapalı çevrelerde, örneğin atari oyunlarında pekiştirmeli öğrenme algoritmaları insan üstü yetenekler sergilemeyi başarmıştır.

Pekiştirmeli öğrenmenin temeli optimal kontrol ve dinamik programlamaya dayanmaktadır. Dinamik programlamanın uğraştığı sorun herhangi bir zamanda belirli durumlar kümesiyle tanımlanmış dinamik bir sistemde sıralı olarak yapılan birden çok kararın istenilen sonuca ulaşacak en verimli şekilde tayin edilmesidir. Bir sistemin dinamik olması verilen her kararın sistemin içinde bulunduğu durumu değiştirmesi anlamına gelir. Art arda verilen kararlar dizisi “politika” (policy) olarak adlandırılmıştır, dinamik programlama verili bir hedefi olan herhangi bir sistemde istenilen sonuca ulaşacak olan optimal politikanın bulunmasıdır. (Bellman, 1954, s: 503-504)

Eğer sistemin geçmiş durumları gelecekte verilecek kararlardan bağımsız ise optimal politika matematiksel bir yöntem olan Markov karar süreciyle hesaplanabilmektedir. Pekiştirmeli öğrenmenin en önemli özelliği ajanın çevre ile etkileşime girmesidir. Ajan çevresiyle etkileşime girdiğinde, yani bir eylemde bulunduğunda söz konusu eylem çevresinin içinde bulunduğu durumu değiştirir. Markov karar sürecinde ajanın her eylemi çevre durumlarının da dönüşmesi anlamına gelir. Ajan her eyleminde aynı dönüşümle çevresinin bilgisini güncellemek durumundadır. Pekiştirmeli öğrenmede belli bir anda yapılan bir eylem bir sonraki anda çevrede bulunan bir ödül ile değerlendirilir. Ödül, ajanın hedeflerine uygun olacak şekilde belirlenmelidir. Ajan toplam zamanda ödülünü maksimize ederek olası optimal politikayı bulmayı hedefler. Deneme yanılma yoluyla ajan belli durumlara uygun eylemi, durum ve eylemin

kombinasyonunun belirlediği ödülü kullanarak ilişkilendirmeyi öğrenir, her durum ve o durumda bulunduğu eylemlerden bir değer tablosu oluşturur. Dolayısıyla pekiştirmeli öğrenme algoritmalarının kullandığı iki ana kavram araştırma ve hafızadır, ajan hedefine ulaşmak için yapacağı eylemleri araştırmalı aynı zamanda belli durumlarda önceden değerlendirilmiş eylemleri kullanabilmelidir. (Sutton, Barto, 1998, s: 61-80)

Bir başka pekiştirmeli öğrenme algoritması olan zamansal fark öğrenmesi ise verilen ödüle göre durumlar ve eylemleri ilişkilendiren bir tablo oluşturmaktan daha çok çevreden gelen bir ödülü tahmin ederken kendi gelecek tahmini de tahmin eder. Bu şekilde ağın öğrenmesi için son ödüle ulaşması gerekmez ve durumlar arasında bir ilişki kurar. Her anda yapılan tahmin dizisi arasındaki fark, zamansal fark hatasını belirler, böylece ajan beklentileri ihlal edildiğinde veya ödüllendiğinde öğrenir. (Sutton, Barto, 1998, s: 133-155) Sutton ve Barto, zamansal fark öğrenmesinin ikinci dereceden bir koşullanma gerçekleştirdiğini göstermiştir. (Sutton, Barto, 1990, s: 497-537) Eğer size beklemediğiniz bir ödül sunulursa orta beyinde dopamin artışı gözlemlenir. Ancak söz konusu ödül her zaman ulaşılabilir kılınırsa dopamin salgılanması durur. Dolayısıyla dopamin ulaşılabilir olduğu kesin ödüller söz konusu olduğunda değil, bir sürpriz durumunda veya ödül ile ilgili bir ipucunun bulunması durumunda salgılanmaktadır. (Schultz, 1998, s: 304-308)

Pekiştirmeli öğrenme tekniklerinde bir başka ayrım ise ajanın baştan içsel bir modele yani bilişsel bir haritaya sahip olması ve ajanın içsel bir modele sahip olmaksızın sıfırdan öğrenerek içsel bir model inşa etmesidir. Model tabanlı öğrenme hesaplama olarak pahalı ve yavaş olmakla beraber çevredeki değişikliklere uyum sağlama yeteneği vardır. Bunun nedeni modelden bağımsız öğrenen bir ajan çevredeki bir değişiklikle eylem ve ödül ilişkisi değiştiğinde tüm değerler tablosunu değiştirmek zorunda kalırken, model tabanlı ajanların çevredeki bir değişiklik durumunda sadece modellerinde ufak değişiklikler yaparak duruma uyum sağlayabilmeleridir.

Pekiştirmeli öğrenmenin derin ağ mimarileriyle birleştirilmesi özellikle satranç, Go gibi kapalı tahta oyunlarında ve atari oyunlarında insan üstü yetenekler sergileyen ajanların tasarlanmasını mümkün kılmıştır. Pekiştirmeli öğrenmenin derin ağlara ilk uygulanması çok katmanlı evrişimli bir ağın girdi verilerini sınıflandırmak yerine girdi durumunun temsillerine göre değer fonksiyonunun değerlerini ilişkilendirecek şekilde tasarlanmasıdır. Bu şekilde ağ hangi durumların temsillerinin ödül, hangi durumların temsillerinin ceza ile ilgili olduğunu yapısındaki ağırlıkları ayarlayarak öğrenmiş ve böylece hedefe yönelik başarılı istatistiki genellemeleri yapmayı başarmıştır. (Mnih,

2015) Aynı mimari dünya Go şampiyonunu yenen AlphaGo'da da kısmen kullanılmıştır. Evrişimli ağlar ile pekiştirmeli öğrenmenin kombine edilmesiyle yaratılan sistemlerin gerçek insan ve hayvan ajanlara kıyasla öğrenmek için çok fazla eğitilme gerek duymaları bu mimarilerin zayıf yönüdür. Ayrıca bu ağların uzun süreli hafızası olmadığında ağın öğrendiklerini unutmaması için öğrenilen durumların sürekli tekrarı gerekmektedir. Bir başka sorun ise ağın belli parametrelerinin önceden ayarlanmasıyla ilgilidir. Eğer ağın ilk ağırlıklarının ayarı ağın hedeflerini belirleyecek şekilde önceden yapılırsa bu ağlar yeni deneyimden öğrenmeyi başaramamakta, ağ hiçbir başlangıç ayarı olmaksızın öğrenmeye başlandığında ise öğrenme süreci çok fazla örneğe ihtiyaç duymaktadır. Atari oyunlarında kullanılan ağlar genel olarak sıfırdan sadece ekrandaki piksellerden ve ödüllerden öğrenmektedir. (Lake, 2017, s: 4)

Bu sorunları çözmek adına getirilen çözümlerden biri mimariye epizodik hafıza eklenmesidir. Bu şekilde ağ her yeni durum vektörünü mimariye eklenmiş bir epizodik hafıza birimine, o durumda uygulamış olduğu eylem ve kazanmış olduğu ödül ile birleştirerek bir değer (Q-değeri) olarak kaydeder. Bu şekilde ağ yeni durumlarda epizodik hafızasındaki vektör durumlarını değerlendirerek yeni durum ile hafızasındaki durum arasında anlık bir ilişki kurmayı başarır. Epizodik hafızanın oluşumu uzun bir süreç olmasına rağmen, ağın deneyimi arttığında, ajan yeni karşılaşılan bir durumda epizodik hafızada saklanan bilgiler seçilerek anlık duruma göre eylemi belirleyecek şekilde kullanılabilmektedir. Bu mimari öğrenme sürecini hatırı sayılır derecede verimli hale getirmiş, hatta öğrenme sürecini insan ve hayvan ajanlarla kıyaslanır seviyelere yükseltmiştir. (Pritzl, 2017, s: 2827) Bu mimari doğal dil işleyen mimarilerdeki dikkat mekanizmalarının epizodik hafızayla birlikte pekiştirmeli öğrenmede kullanılmasına tekabül eder.

Derin zamansal fark algoritmalarında devrim yaratan diğer bir teknik ise meta-öğrenme, bir başka ifade ile öğrenmeyi öğrenmek olmuştur. Meta-öğrenme bilişsel bilimlerde ortaya atılmış bir kavramdır. 1949 yılında Harlow maymunlarla yaptığı deneyde maymunun önüne iki farklı nesne koymuş ve maymunun nesnelerden birini seçmesini istemiştir. Nesnelerin birinin altında ödül bulunmaktadır. Maymun ilk nesneyi seçtikten sonra aynı iki farklı nesne altı kez daha maymunun önüne konmuş ve ödül hep aynı nesnenin altına saklanmıştır. Her tekrarda tamamen yeni farklı nesneler kullanılmış ancak az sayıda tekrardan sonra maymun her deneyde ödülün sadece nesnelerin biriyle ilişkili olduğunu öğrenmiştir. Aynı deney insan çocuklarıyla

da tekrarlanmış ve benzer sonuçlar elde edilmiştir. Bir başka deyişle maymunlar ve çocuklar öğrendikleri ilişkiyi daha önce görmedikleri nesnelere uygulamayı başaracak basit bir kuralın genellemesini yapmayı başarmaktadırlar. Bu genelleme yeteneği endüktif bias olarak tanımlanır. (Harlow, 1949, s: 51-65) Seri bilgiyi işleyecek şekilde tasarlanmış tekrarlayan sinir ağlarında çıktı zamansal fark öğrenmesiyle durum, eylem ve ödül ile ilişkilendirildiğinde ağın Harlow deneyinde maymun ve çocuklarla aynı performansı sergilediği gözlemlenmiştir. Bu tip mimarilerde tekrarlayan sinir ağlarında oluşan çalışan belleğe zamansal fark öğrenmesi ile belli aktivasyon desenlerinin de öğretilebileceği anlaşılmıştır. Bu noktada zamansal fark öğrenmesinin ağın ağırlıklarını ayarlamaktan (ağın ağırlıkları dış dünyanın temsilleriyle güncellenir) daha çok aktivasyon desenlerinin yapısı üstünde öğrenmesi genellemeyi mümkün kılmaktadır. (Wang, 2018, s: 860-868)

Derin pekiştirmeli öğrenmeyi biyolojik öğrenmeden ayıran en önemli iki farktan ilki bu tip ağların öğrenmeye sıfır noktasından deneyimle başlamasıdır. Ancak biyolojik canlılar öğrenmeye doğuştan sahip oldukları birçok bilgi ile başlar. Bir diğer fark ise biyolojik canlıların değer fonksiyonlarının evrimle gelişmiş duygu mekanizmaları olması, derin pekiştirmeli ağlarda ise değer fonksiyonlarının ağın hedeflerini belirlerken tayin edilmesidir.

Derin sinir ağları henüz insanın zihinsel yetilerini taklit etmekten çok uzak olsa dahi bu ağlar beynin basit duyusal yetilerini taklit edebilmektedir. Epizodik bellek, çalışan hafıza gibi kavramların bilinçle alakalı kavramlar olması da özellikle derin pekiştirmeli öğrenmenin canlı zihinlerini taklit edebileceği düşüncesini desteklemektedir. Ancak dünyada soyutlamalardan oluşan dil kullanımı sadece insana mahsus bir yetenektir ve derin mimariler henüz insan zihnini taklit etmek için gerekli olan birçok özellikten yoksundur. Derin ağların insan zihnini taklit edebilmesi için beynin yapısındaki talamus, hipokampüs gibi alt sistemlerin korteksteki sinir ağlarıyla fonksiyonel iletişiminin mekanizmalarının anlaşılması elzem görünmektedir.

## 2.7 BAĞLANTICILIK VE SEMBOLİK SİSTEMLER

Yapay zeka çalışmalarının birbiriyle rakip iki odağı olan sembolik sistemler ve bağlantıcılık temelli modeller kıyaslandığında her iki yaklaşımın da güçlü ve zayıf yönleri olduğu ortaya çıkar. Günümüzde güçlü yapay zekanın bağlantıcılık temelli modeller ile daha mümkün olduğu görüşü ağırlık kazanmış olsa da bağlantıcılık

temelli modellerin insan zihninin birçok yetisini taklit edemeyeceğini savunan görüşler bulunmaktadır.

Sembolik sistemleri savunan ve insan zihninin yapay sinir ağlarıyla taklit edilemeyeceğini belirterek bağlantıcılığı eleştiren görüşlerin temelinde zihinsel durumlarının, özellikle de insan dilinin ancak sembolik bir yapıyla açıklanabileceği düşüncesi yatmaktadır. Dennett, felsefe dünyasının insan zihnine bakışını inceleyen, "The Logical Geography of Computational Approaches: A view from the East Pole" makalesinde hesaplamalı zihin kuramı içindeki fikirleri coğrafi bir metaforla incelemiş, haritanın merkezinin doğu kutbu olduğunu, doğu kutbunun merkezinde ise MIT'nin olduğunu belirtmiştir. Dennett'ın metaforik haritasında Doğu kutbundan uzaklaşan her fikir batıya doğru ilerler. (Dennett, 1986, s: 62) Doğu kutbunun temelinde ise Chomsky'nin dil kuramı ve bu kuramla bağlantılı olan Fodor'un düşünce dili hipotezi bulunmaktadır.

İnsan dilinin temelindeki mekanizmaların anlaşılması gerektiğini, daha sonra bu mekanizmaların öğrenme algoritmalarıyla nasıl edinildiğini ve dili kullanırken algoritmaların nasıl işlediğinin anlaşılması gerektiğini ilk olarak Chomsky tarafından öne sürmüştür. (Chomsky, 1965, s: 5) Chomsky insan dilinin temelindeki iki özeğe dikkat çekmiştir. İlk olarak insan zihni sınırlı sayıda kelimeler kullanarak sınırsız sayıda cümleler kurma yetisine sahiptir. Bir başka ifade ile insan zihni kelime, cümle gibi sembol yapılarını oluşturma kabiliyetine, ardından bu sembolik yapıların iyi biçimlendirilmiş olmasına olanak veren dilbilgisi kurallarını kullanma yetisine sahiptir. Bu yeti ise ancak insan zihninin belli zihinsel dilbilgisi kurallarına sahip olmasıyla açıklanabilir. Dilin bir başka dikkat çekici özelliği ise çocukların bu dilbilgisi kurallarını bir eğitime gerek kalmaksızın doğal olarak öğrenebilme kabiliyetleridir. Chomsky, çocukların doğuştan dilbilgisi kurallarını öğrenme kabiliyetine sahip olduğunu iddia etmiştir. (Pinker, 2018, s: 19) Dilin sınırlı araçlarla sınırsız kullanım oluşturmasını sağlayan kurallar bütünü üretici dilbilgisi olarak tanımlanmıştır. (Pinker, 2018, s: 92) Chomsky dili anlamlı yapan dilbilgisi kurallarını incelemiş ve dünya üstündeki tüm dillerin belli ortak kurallarına sahip olduğunu belirtmiştir. Bu kurallar tüm insan dillerinde ortak ve insanların doğuştan sahip oldukları kurallardır. (Pinker, 2018, s: 123) "Chomsky'e göre, ziyaretimize gelecek Marslı bir bilim adamı, karşılıklı anlaşılmaz kelime hazineleri hariç, Dünyalıların tek dil konuştuğu sonucuna varacaktır." (Pinker, 2018, s: 261-262) Dolayısıyla bütün dünya dillerinin altında simgeleri yönlendiren aynı kurallar dizisi bulunmaktadır. Bu kurallar dizisi dili mümkün

kılar. “Üretme ve anlama arasındaki doğal ve ortak bir dilbilgisi şifresi, konuşulanların anlaşılır dilbilgisel mesajlar üretmelerine ve bunun tersini yapmalarına imkan sağlar.” (Pinker, 2018, s: 268)

İnsan zihninin sembolik yapılarla işlediğini savunan bir başka filozof olan Fodor, “düşüncenin dili” teorisinde, farklı lisanlar konuşan insanların aynı kavram ve fikirlerden bahsettiğini, dilin temelinde zihinsel bir düşünce yapısı olduğunu ve söz konusu yapının konuşulan dilin yapısıyla aynı yapıya sahip olması gerektiğini iddia etmiştir. “Fodor’a göre, dinleyenin önyargılarıyla ve beklentileriyle bozulmadan, konuşmanın tümcesini kelimesi kelimesine ileten bir tümce algılama modülü, her yerde ve her zaman aynı olan, evrensel olarak yapılanmış bir insan zihni olduğunun simgesidir, bu da insanların, tercih, gelenek ve şahsi çıkardan bağımsız olarak gerçeklik adına, doğru ve gerçek olanın üzerinde uzmanlaşmalarına imkan verir.” (Pinker, 2018, s: 462) Dilin öğrenilmesi ise ancak doğuştan bir mekanizmaya sahip olmakla mümkündür, eğer öğrenmeyi mümkün kılan bir mekanizma söz konusu değilse öğrenmenin gerçekleşmesi mümkün değildir. Bir başka ifadeyle nasıl ki dilbilgisi evrensel bir tasarımsa bu tasarımı kullanılabilir yapan semboller üstünden işlev gören evrensel bir zihin tasarımı da olmak zorundadır. (Pinker, 2018, s: 467)

Ancak bilişsel bilimler alanında Rosch, Chomsky’nin dilsel yeterlilik kuramına karşı çıkarak, insan zihnindeki kategorilerin formel tanımlar ve kurallardan oluşmadığını, kategorilerin genel olarak algılanan örnekler üstünden yaratıldığını öne sürmüştür. Klasik sınıflandırma kuramında bir kategorinin her üyesi o kategorinin özelliklerini taşımak zorundadır dolayısıyla bir kategoriyi temsil eden bir örneğin, o kategorinin başka bir örneğinden daha bir örnek olması söz konusu değildir ve kategorideki tüm elemanların birbirine benzemesi gerekir. Rosch, gerçek yaşamda bir kategorinin bazı üyelerinin, o kategoriyi, kategorinin diğer üyelerinden daha iyi temsil ettiğini öne sürer. Kategori hiyerarşilerini dikey ve yatay boyut olmak üzere iki boyutta tanımlayan Rosch, dikey düzlemin kategorinin kavramsal kapsamlarını temsil ettiğini, yatay düzlemde ise kavramsal kapsamların örneklerinin bulunduğunu belirtir. Ancak, hiyerarşinin en üst seviyesinde o kategoriye ait örneklerden zihinsel bir imge yaratılamaz, üst seviyede kategorinin temel özeliklerinin kavramsal betimlemeleri bulunur ve söz konusu olan örneğin o kategoriye girebilmesi için üst seviyedeki kavramsal koşullara uygun olması gerekir. (Rosch, 1978, s: 29-30) Zihinsel bir imge ile örneklenebilen ilk dikey seviye orta kısımlarda bulunan temel seviyedir. Eşya, sandalye, çalışma sandalyesi örneğini veren Rosch, eşya kategorisinin genel

özelliklerine sahip ama nötr bir zihinsel imge yaratılmayacağını ancak temel seviyede zihinsel sandalye imgelerinin düşünülebileceğini belirtir. (Rosch, 1978, s: 32) Temel seviyede bir sandalye düşünüldüğünde ise yatay hiyerarşide, belli sandalye imgelerinin kategoriyi diğer sandalye imgelerinden daha iyi örneklendirdiğini belirten Rosch, bu örneklere kategorinin prototip demiştir. Prototipler bir anlamda o kategorinin en genel ve yaygın örnekleridir. (Rosch, 1978, s: 35) Ayrıca Rosch, temel seviyedeki örneklerin çocuklar tarafından ilk algılanan örnekler olduğunu ve çocukların kavramları temel seviye örnekleri ile öğrendiğini göstermiştir. Yapılan deneyler prototip örneklerin ayrıcalıklı olduğu, prototip örneklerin denekler tarafından daha hızlı algılandığını ve deneklerin prototip örnekler üstünden daha hızlı ve daha kesin muhakemeler yapabildiklerini göstermiştir. (Rosch, 1978, s: 38)

### 2.7.1 Semboller ve Alt-semboller

Rosch'un kuramı insan zihninin en azından temel düzeyde sembol ve sembol manipülasyonlarıyla işlemediğini, zihnin temel seviyede istatistiki sınıflandırmalar yaparak soyut düzlemde kategorileri ve kavramları oluşturduğunu savunmuştur. Smolensky zihnin işleyişinin üç temel seviyede anlaşılabileceğini öne sürmüştür. Söz konusu üç seviye nöronal seviye, alt-sembolik seviye ve sembolik seviye olarak tanımlanabilir. Alt-sembolik seviyedeki ince yapılı parçalar, bir üst seviye olan sembolik seviyede çalışan sembol ve sembol manipülasyonlarının yapıtaşlarını oluşturur, alt-sembolik seviye ile sembolik seviye arasındaki ilişkiyi Newton fiziği ile kuantum fiziği arasındaki ilişkiye benzetmek mümkündür. (Smolensky, 1988, s: 12)

Smolensky zihnin seviyelerinde kavramsal ve alt-kavramsal olmak üzere yaptığı ayrımda alt-kavramsal seviyeyi alt-sembolik seviye, kavramsal seviyeyi ise sembolik seviye olarak tanımlar. Sembolik yapıda kendini gösteren ve sembol manipülasyonlarının gerçekleştiği kavramsal seviyede iş gören "bilinçli kural yorumlayıcısı" kültürel bilginin ajanlar arasında paylaşımını sağlayan sanal bir makine olarak düşünülebilir. Kültürel bilgi aynı dili kullanan tüm ajanlar tarafından paylaşılabilir, farklı ajanlar sembol manipülasyonlarıyla türetilmiş olan sonuçları teyit edebilir ve aynı zamanda sembollerin kullanıldığı çıkarım işlemleri tecrübesi olmayan ajanlar tarafından kolayca öğrenilebilir. Dolayısıyla sembolik seviye zihnin en üst seviyesidir, bilinçli süreçler sembolik seviyede meydana gelirler. Ancak bu seviyenin altındaki alt-sembolik seviye bir sanal makine yapısına sahip olmaktan daha çok

nöronal seviyeye daha yakındır. Alt sembolik seviye nöronlar ve nöronların aralarında kurduğu dinamik ilişkileriyle ifade edilebilir. Alt-sembolik seviyedeki ince yapılı dağınık dağıtılmış nöronal yapılar ise kavramların yapıtaşlarıdır. (Smolensky, 1988, s: 3)

Çok boyutlu bağlantı uzayında bir grup nöronun ifade ettiği temsil, tüm nöronlar arasındaki bağlantı ağırlıklarından oluşan çok boyutlu bir vektör ile ifade edilebilir. Söz konusu vektör ise, örneğin söz konusu nöron kümesi bir nesneyi temsil ediyorsa, nesneyi tanımlayan sembole denk gelir. Dolayısıyla sembol dönüşüm kuralları ve kavramlar alt-sembolik seviyedeki nöron örüntülerinin ağırlıklarıyla tanımlanan vektörlere denk gelmektedir. Bir sembol ifade ettiği kavramdan bağımsız iken alt-sembolik temsillerden oluşturulan vektörler ifade ettikleri semboller arasında belli bir benzerlik ve farklılık ilişkisini ifade ederler. Bu durum beyinde benzer nesnelerin nöron temsillerinin birbirileriyle ortak nöron devreleri kullanmasıyla ilişkilidir. (Smolensky, 1988, s: 11) Smolensky alt-sembolik seviyede bilginin, sembolik seviyedeki gibi açık bir temsili olmadığını ancak bu bilginin belli bir şekilde ifade edildiğinde sembolleri ve semboller arasındaki kuralları oluşturabileceğini savunmuştur.

Sembolik seviyedeki hesaplamalar sembollere seri şekilde uygulanan kuralları içerirken, alt-sembolik seviyedeki hesaplar paralel ve sayısaldır. Alt-sembolik seviyeyi nöron kümeleri ve söz konusu kümedeki nöronlar arasındaki ağırlıklarla tanımlanır. Alt-sembolik yapı tamamen sayısal olarak belirlenmiştir. Bu seviyede özellikle tekrarlayan sinir ağları türevindeki dinamik sinir ağlarının girdiye ait bilgiyi, bilgideki benzerlik ve farklılıklara göre çok boyutlu uzayda temsil eder. Ağın taneli yapısında kümeler oluşturmuş her bir temsil alt-sembolleri oluşturur. Alt semboller ise sembollerin yapı taşlarıdır. Ayrıca nöron kümeleri belirli düzenlilikleri yani sembol dönüşüm kurallarını da temsil etme yetisine sahiptir. Tekrarlayan sinir ağlarının sadece girdi verilerinden yola çıkarak dilbilgisi kurallarını hatasız olarak öğrenebildiği gerçeğinin ortaya çıkması bu görüşü desteklemiştir.

### 2.7.2 Doğal dil işleme

Doğal dil işleme için kullanılan tekrarlayan sinir ağları, aynen ses işlemede kullanılan tekrarlayan sinir ağları gibi, belli bir dizi girdisinden sonra hangi terimin geleceğini tahmin ederek iş görürler. Doğal dil işlemede veya ses işlemede kullanılacak yapay sinir ağının tasarımındaki en büyük teknik sorun, birbiriyle bağlantılı bir dizinin ağa nasıl gösterileceğinin bulunması olmuştur. En nihayetinde ses dizileri ya da doğal dil

zaman içinde art arda gelen bilgi parçalarının birleştirilmesini gerektirir. Yapay sinir ağı araştırmacılarının bu soruna bulduğu çözüm doğrusal bir ağın son birimindeki çıktıyı ağın ilk birimine vererek tekrarlayan bir sinir ağı oluşturmak olmuştur. (McClelland, 2008, s: 31) Bu şekilde ağa bir hafıza verilmiş olur ve ağ son girdiyle aynı anda bir önceki girdiyi birleştirir. Ağa art arda belli sayıda kelime verildiğinde ise en son adımda ağ, tekrarlayan bağlantılar sayesinde, tüm kelimelerin birleştirildiği bir temsile ulaşmış olur. (Elman, 1990, s: 181-185) Bu tip ağların eğitimi ise zamana bağlı geri yayılım ile yapılabilmektedir. Bu teknikte her girdi verisi bir zaman biriminde ağa verilmiş gibi düşünülerek geri yayılım zamanda geri giderek ağı eğitir. Ağ eğitildikten sonra her dizinin devamını tahmin etmeye çalışarak iş görür.

Doğal dil işlemede tekrarlayan sinir ağının girdi dizisi uzadıkça ilk girdilerin ağırlıklarının azalması, yani ağın ilk girdilerden daha çok son girdilere göre hareket etmesinin ağlarının başarısını düşürdüğü anlaşılmıştır. Yapay zeka araştırmacılarının hafıza sorununa çözümü ise Hopfield ağlarındaki gibi bir ilişkisel belleği ağa vermek olmuştur. Uzun-kısa bellek birimleri birbirileriyle bağlantılı üç kapıya karşılık gelen üç derin ağdan oluşturulmuş tekrarlayan sinir ağlarıdır. Okuma kapısı son girdiyi alır, diğer iki kapıdan sil kapısı dizdeki yeni terime göre önceki bilgilerden hangilerinin unutacağına karar verir, yaz kapısı dizinin son halini bir düşünce vektörü olarak ifade edilmesini sağlar. Bu şekilde ağ birçok kelimeden oluşan bir cümleyi tek bir vektör yapısında temsil etmeyi başarır. (Hochreiter, 1997, s: 1740-1744)

Bir dilden başka bir dile çeviri yapmanın en büyük zorluğu cümle yapılarının her dilde farklı olabilmesidir. Tekrarlayan sinir ağlarından önce makine çevirisi iki adımda yapılmaktaydı. İlk adımda derin bir sinir ağı dilin içerdiği kelimeleri birbiriyle ilişkilerine göre sayısal vektörlere çevirirdi. Bu yöntemde ağ belli bir yazılı metindeki kelimelerin hepsinin diğer kelimelerle ilişkisini de kale alarak  temsil edecek şekilde eğitilirdi. Bu şekilde ağın yüksek boyuttaki temsilinde birbirleriyle yakın anlam taşıyan tüm kelimeler belli gruplar oluştururdu. İkinci adımsa iki farklı dilin gramer yapılarına göre çevirinin kuralları el ile belirlenirdi. (LeCun, 2015, s: 440)

Tekrarlayan sinir ağlarında ise makine çevirisi yapılırken kullanılan ağ iki dilin dilbilgisi kurallarını ve iki dil arasındaki çeviri ilişkisini başka bir veriye gerek kalmaksızın kendi başına öğrenebilmektedir. Makine çevirisi yapan sistemlerin mimarisi birbiriyle tek bir ağ oluşturacak şekilde bağlanmış iki tekrarlayan sinir ağından oluşur. İlk tekrarlayan sinir ağı çevrilecek cümle veya metni kodlayarak tek bir düşünce vektörü oluşturur. İkinci tekrarlayan sinir ağı ise ilk ağın kodladığı düşünce vektörüne bağlı kalarak

istenilen dilde cümle veya metin üretir. Bu mimaride ilk tekrarlayan sinir ağına kodlayıcı, çevirinin sonucunu üreten tekrarlayan sinir ağına ise çözümleyici denmektedir. Ağın eğitimi iki dilde de var olan paralel bir metin üstünden yapılır ve tüm sistem için tek bir zamanda geri yayılım kullanılarak eğitilir. (Bahdanau, 2015, s: 2-4)

Bu mimari ile özellikle kısa cümlelerin çevirisinde büyük başarı sağlanmıştır. Ancak makine çevirisinde kodlayıcı tüm cümlenin içindeki veriyi tek bir vektörde topladığından özellikle bağlaçların kullanıldığı uzun cümlelerin çevirisinde kısa cümlelerin çevirisindeki başarıyı gösterememiştir. Yapay sinir ağı araştırmacılarının uzun cümlelerin çevirisindeki başarıyı yükseltmek adına kullandıkları dikkat algoritması kullanmışlardır. Dikkat algoritması beyindeki dikkat mekanizmalarından esinlenmiş bir tekniktir. Dikkat mekanizmasında çözümleyici tekrarlayan sinir ağının çeviri cümlesini üretirken kodlayıcı ağın kodlamayı yaparken kullandığı her kelimeye bakması yani her kelimedeki bilginin son cümlede kapsandığını kontrol etmesine olanak sağlayan geriye dönüşlü bağlantılar kullanılır. (Hassabis, 2017, s: 247) Bu şekilde uzun cümlelerde meydana gelen bilgi kaybı mümkün olduğunca azaltılması hedeflenir. Dikkat mekanizmaları uzun cümlelerin çeviri başarısını yükseltmiştirler.

### 2.7.3 Zihnin Sistematikliği

Doğal dil işlemede kullanılan tekrarlayan sinir ağlarının dilbilgisi kurallarını içsel yapılarında oluşturma yetisine sahip olduğunun gösterilmesi, yapay sinir ağlarının açık sembolik yapıları da oluşturabileceği düşüncesini desteklemiştir. Ancak Fodor ve Pylyshyn bağlantıcılık temelli modellerin sadece verinin istatistiki derlemesi sayesinde dağıtılmış temsiller ve bu temsillere dair bir hafıza oluşturabildiğini belirtmiştir. Ancak bağlantıcılık temelli modellerin sembol sistemleri tarafından oluşturulabilen ve düşüncenin sahip olduğu üretkenlik, sistematiklik, birleşimsellik ve çıkarımsal tutarlılık özelliklerini yapılarında bulundurmalarının mümkün olmadığını savunmuştur. Söz konusu dört özellik birbiriyle bağlantılı özelliklerdir. (Fodor, Pylyshyn, 1988, s: 21-33)

Üretkenlik insan zihninin, insan dilinde olduğu gibi sınırlı kaynaklar kullanarak sınırsız sayıda temsil üretme yetisidir. Bu yeti insan zihninin yeni konularla başa çıkma becerisinin bir paçasıdır. Bir insan daha önce hiç karşılaşmadığı bir durumda nasıl davranacağını içsel temsillerinin üretkenliğiyle tayin eder. İnsan düşüncesinin sistematikliği insan zihninin belirli düşüncelerden yola çıkarak kurduğu bağlantılarla

verili olmayan bir düşünceyi yaratma yetisidir, örneğin John Mary'i seviyor düşüncesinden yola çıkarak Mary'nin de John'u sevdiği düşüncesinin düşünebilir olması gibi. (Fodor, Pylyshyn, 1988, s: 17-18) Sistematiklik aynı zamanda birleşimsellik ile ilişkilidir. Birleşimsellik zihin temsillerinin birleşen kısımlarından oluştuğunu öne sürmüştür. Komplike temsiller ve anlamları kısımların yeniden düzenlemesiyle oluşturulur. Aynı şekilde çıkarımsal tutarlılık bir önermedeki ilişkileri, örneğin P&Q önermesinden P ve Q önermelerini ayırabilme yetisidir. (Fodor, Pylyshyn, 1988, s: 32) Fodor düşüncenin sahip olduğu üretkenlik, sistematiklik, birleşimsellik ve çıkarımsal tutarlılık özelliklerinin ancak sembolik bir sistemde sentaksın belirlenmesiyle mümkün olduğunu savunmuştur. Fodor, bilişsel sistemlerin temsillerinin seviyelerinden bağlantıcılık temelli modellerin ancak dünyaya ait temsilleri alt seviyede içsel temsillerinde oluşturabildiğini ancak bilişin üst seviye özelliklerini temsil edemediklerini vurgulamıştır. Günümüzdeki derin ağlar sistematiklikten yoksundur, Fodor'un da belirttiği gibi dinamik yapay sinir ağları dahil tüm yapay sinir ağları temelde gelişmiş istatistiki çıkarım makineleridir ve iki farklı olay arasında sistematik bir bağlantı kurma yetisine, bir başka ifadeyle belirli bir dağılımın ötesinde genelleme yetisine sahip değildirler. Genelleme yetisi iki farklı nöronal temsil arasında hiç kurulmamış bir ilişki kurarak bir nöronal temsilin nöron bağlantı ağırlıklarına içkin kuralları tamamen farklı bir nöron örüntüsüne uygulanmasını içermek zorundadır. Sembolik sistemlerse sembollerin temsillerini ifade etmeyi başaramamakta, bir başka ifadeyle sembolleri ile fiziksel anlamları arasında bir ilişki kuramamaktadır.

Sonuç olarak hem sembolik sistemlerin hem de bağlantıcılık temelli modellerin çözmesi gereken ciddi sorunları bulunmaktadır. Bağlantıcılık temelli modellerde sistemin iç temsillerini yaratabiliyor olması bu modellerin ileride başarılı olabileceği düşüncesini desteklemektedir. Pekiştirmeli öğrenmenin baglantıcılık temelli modeller ile uyumlu bir şekilde birleştirilebilmiş olması da bu düşünceyi desteklemektedir. Bağlantıcılık temelli modeller beynin yapısından ilham almaktadır ki nörobilim araştırmalarında epizodik bellek oluşumu veya dil yetisinin nasıl işlediği anlaşıldığında bağlantıcılık temelli modellerin başarısının artacağı düşünülebilir. Bağlantıcılık temelli modellere olan inancın temeli ise beynin işlevlerini belli bir organizasyona sahip nöronlar tarafından gerçekleştiriyor olmasıdır.

# 3. BÖLÜM

# YAPAY ZEKANIN OLANAKLILIĞINA DAİR FELSEFİ İTİRAZLAR

Yapay zeka teknolojisinin öncüleri araştırmalarını ileride makinelerin sadece hesap yapmaktan öteye gidip düşünebileceği hatta bilinçli olabileceğine inanarak yönlendirmiştir. Marvin Minsky yapay zekayı, insan zekasını gerektiren işleri makinelere yaptırma bilimi olarak tanımlarken, Haugeland, Minsky'nin düşüncesini bir adım ileri taşıyarak yapay zeka çalışmalarının temel amacının, zekayı taklit etmek ya da zekice bir kopya üretmek olmayıp, tam ve gerçek anlamda zihni olan makinler tasarlamak olduğunu öne sürmüştür. Weizenbaum da aynı görüşe eşlik etmiş ve yapay zeka araştırmalarının her şeyden önemli olan amacının bir insan zekasına eşit, hatta insan zekasından üstün yapay sistemler yaratmak olduğunu vurgulamıştır. (Dreyfus, 1972, s: xv-xxviii)

Yapay zeka araştırmacılarının çoğu insan zekasına eşit, hatta insan zekasından üstün yapay sistemler yaratılabileceği inancı doğrultusunda ilgilerini insan zihninin modellenmesi üstüne yoğunlaştırmış ve tartışmalarını sembolik mantık ve bağlantıcılık çerçevelerinde sürdürmüşlerdir. Ancak tüm yapay zeka araştırmacıları bilgisayarların düşünebilen, bilinçli makineler olacağı inancını paylaşmamış, çoğu, bilgisayarların sadece karışık hesaplar yapabilen zeki sistemlerden öte bir noktaya ulaşamayacağını savunmuştur. Bu fikir ayrılığı, "zayıf yapay zeka" ve "güçlü yapay zeka" kavramlarını üretmiştir. İlk kez Searle tarafından öne sürülen bu kavramlardan zayıf yapay zeka, çok karmaşık problemleri çözebilen ama temelde hesap yapmaktan öteye gitmeyen sistemler için kullanılırken, güçlü yapay zeka kavramı öğrenebilen, düşünebilen, kendi farkındalığı olması hedeflenen sistemleri betimlemek için kullanılmaya başlanmıştır.

## 3.1 SEARLE'ÜN SENTAKS VE SEMANTİK VURGUSU

Searle'ün düşüncesine göre bilgisayarlar çok karışık hesapları yapabilmenin ötesine geçemezler. Searle, doğru girdilere ve doğru çıktılara sahip olan doğru şekilde

programlanmış dijital bir bilgisayarın insanların zihne sahip olduğu gibi bir zihne sahip olamayacağını belirtmiş ve yapay zekayı, “doğru girdilere doğru çıktılar üretme kapasitesi olan bir dijital bilgisayar” olarak tanımlamıştır. (Searle, 1980, s: 417) Bu şekilde tasarlanan bir bilgisayarın Turing testini geçeceği varsayılır. Ancak Searle bu şekilde programlanan bir bilgisayarın belli sembol manipülasyonlarının ancak belli kurallarını uygulayabileceğini ancak asla bir anlayışa sahip olamayacağını da iddia etmiş, güçlü yapay zekanın olanaksızlığını göstermek amacıyla “Çince odası” düşünce deneyini tasarlamıştır.

Çince odası düşünce deneyinde Searle bir odadadır, odanın içinde Çince yazılmış mesajları yanıtlama talimatlarını içeren bir kural kitabı bulunmaktadır. Searle, sadece Çince karakterleri tanıyıp kural kitabındaki talimatları uygulayarak gelen mesajları cevaplandırır. Searle’ün iddiası eğer kural kitabının talimatları yeterince başarılı yazılmışsa gelen mesajlara doğru cevaplar vereceğini ancak Çince bilmediğinden mesajlaşmanın içeriğini anlayamayacağıdır. Searle’ün deneyinde, Searle’ün kendisi, kural kitabı, girdi çıktı sistemleri bir bilgisayarı betimlemektedir. Searle, doğru girdilere doğru cevaplar üretmiş ancak kendisi girdilerin ve çıktıların anlamını anlamamıştır.

Searle, Turing testini geçen bir bilgisayarın anlama veya düşünme yetisi olamayacağını Çince odası düşünce deneyiyle kanıtladığını öne sürer. Bir dili anlamak için belli sembollere dilbilgisi kurallarını uygulamak yeterli değildir, semantik belli zihinsel temsillerin varlığını gerektirir. Searle, bilgisayar programlarının zihinsel temsillere sahip olmaksızın belli kuralları takip eden sistemler olduğunu vurgular. Bir başka deyişle bilgisayar programları sentaktiktir, insan zihinleri ise zihinsel içeriğe, dolayısıyla semantiğe ve anlama yetisine sahiptir. Ancak Searle, Çince odası deneyinde Minsky’nin “Yapay Zekaya Doğru” makalesinde belirttiği fikri hesaba katmayarak düşünce sürecinin sadece belli sembol manipülasyonlarından oluştuğunu varsaymıştır. Minsky’nin ortaya koymuş olduğu fikir ise düşünebilen bir makinenin kendi iç temsilleri olması gerektiğidir.

Searle’ün düşünce deneyine birçok yanıt verilmiştir. Bu yanıtlardan biri “sistem yanıtı”dır. Sistem yanıtına göre Searle’ın mesajları anlamaması normaldir, çünkü Searle sistemin sadece bir parçasıdır. Ancak sistem bütün olarak düşünüldüğünde kural kitabı bilgi bankası, Searle işlemci olmakta ve girdi ve çıktı birimleriyle birlikte tüm sistem Çince anlamaktadır. (Searle, 1980, s: 419)

Çince odasına verilen içsel temsillerin zorunluluğu ile ilgili olan yanıt ise “robot yanıtı”dır. Bir odaya kapatılmış bir robotun Çince kelimelerin anlamlı olduğunu dahi anlaması imkansızdır. Çin odası deneyindeki mesajların semantik anlam taşıması ancak sistemin dış dünya temsillerine ulaşıp, dış dünya temsilleri ile yazılan mesajlar arasında nedenselliğe dayanan bir ilişki kurmasıyla mümkün olabilir. Robot yanıtı, güçlü yapay zekanın sadece belli kurallar dizisiyle oluşturulamayacağını, anlamın değişik temsillerin birbirileriyle nedensellik ilişkileri içinde etkileşmesiyle mümkün olabileceğini iddia etmektedir. Searle’ün robot yanıtına cevabı, kendisinin, robotun beynini temsil eden bir odada kural kitabıyla mesajları cevapladığı taktirde, yine de mesajları anlamayacağıdır. (Searle, 1980, s: 421) Searle’ün bu cevabı dış dünya temsillerinin robotun devrelerine dağıtılmış olmaktansa merkezi bir yerde olduğunu varsaymıştır.

Searle’ün Çince odası düşünce deneyi esasen sembolik mantık ve kural dizileriyle çalışan eski yapay zeka sistemleri düşünülerek tasarlanmış bir düşünce deneyidir. Eski yapay zeka sistemlerinin temelinde ise fiziksel simge hipotezi bulunmaktadır, fiziksel simge hipotezi sembol yapılarının dönüştürülmesiyle çalışan sistemlerin zeki eylemler için gerekli ve yeterli imkanlara sahip olduğunu öne sürer. Ancak Searle’ün örneğinde ya da eski yapay zeka sistemlerinde kullanılan sembollerin fiziksel dünya ile ilişkisinin kurulamamış olması fiziksel simgelerin temellendirme sorunu olarak bilinir.

Ancak Searle, her türlü yapay zeka sistemi için Çince odası düşünce deneyinin geçerli olacağını belirtir. Searle’e göre hesaba dayalı zihin modelleri, zihinle ilgili birçok kavramı dışarıda bırakmaktadır. Searle’ın semantik eleştirisi günümüzdeki yapay sinir ağları açısından değerlendirildiğinde tekrarlayan sinir ağlarının semantiğe sahip olmadığı açıktır. Ancak özellikle tekrarlayan sinir ağlarının kendi temsillerini çevre ile etkileşim içinde yani dinamik olarak oluşturup, güncelleyebiliyor olması yapay sinir ağlarının ileride semantiğe sahip olacak şekilde tasarlanabileceği umudunun nedenidir. İnsan beyninde semantiğin nasıl oluştuğuna dair modellerin ortak noktasıysa dil yetisini oluşturan sinir ağının tüm algı ve motor hareket modüllerini kapsayacak şekilde bu modüller ile  bağlantılı olduğunu kabul etmeleridir.

## 3.2 FENOMENOLOJİK İTİRAZ

### 3.2.1 Bilgisayarlar Ne Yapamaz

Hubert Dreyfus, güçlü yapay zekanın olanaklı olmadığını ilk savunan filozoflardandır. Dreyfus, insan zekasının makineler tarafından taklit edilmesinin mümkün olmadığı görüşünü fenomenolojiyle temellendirerek savunmuştur. Hubert Dreyfus, 1972 yılında yayınladığı, “Bilgisayarlar Ne Yapamaz” kitabında insana özgü bilişsel yetilerin simgesel kodlar ile neden gösterilemeyeceğini açıklamıştır.

Dreyfus eleştirilerini simge hipotezine dayanan yapay zeka mimarilerini inceleyerek temellendirmiştir. (Dreyfus, 1972, s: 67) Dreyfus, bu mimarilerin dört temel varsayıma dayandığını belirtir. İlk varsayım olan biyolojik varsayıma göre beyin işlevsel seviyede bilgiyi ayrık işlemlerle işler. İkinci varsayım psikolojik varsayımdır; bu varsayıma göre zihin formel kurallar ile çalışmaktadır. Üçüncü varsayım ise tüm bilginin semboller ile formüle edilebileceğini belirten bilgi varsayımıdır. Dördüncü varsayım ise dış dünyayı betimleyen temsillerin birbirilerinden bağımsız olduğunu belirten ontolojik varsayımdır. (Dreyfus, 1972, s: 68)

Dreyfus, Heidegger felsefesinden yola çıkarak, dünyada tek başına anlam taşıyan varlıklar olmadığını, varlıkların anlamlarını diğer varlıklarla girdikleri ilişkilerin belirlediğini ve anlamın her zaman teknik bilgiden fazlası olduğunu vurgulamıştır. Bir varlığa ait anlamlı bir iç temsilin oluşması ancak günlük hayattaki varlığa ait sonsuz sayıdaki anlamların bütünü ile mümkündür. Ayrıca temsil edilen her varlığın, diğer varlıklarla alakasının da belirlenmesi gerekmektedir. İnsan, dünyaya fırlatılmış bir varlık olarak belirli bir durumda belli bir nesnenin hangi ilişkilerinin önemli, hangilerinin ise önemsiz olduğunu deneyimleriyle bilirken, simge hipotezi ile nesnenin hangi ilişkilerinin, hangi durum için öncelikli olduğunun çıkarılması mümkün değildir. (Dreyfus, 1972, s: 177-179, s: 188-189)

Minksy, yeni bir durumda sistemin, o duruma ait çerçeveyi (durumla ilgili tüm ilişkilerin tanımlandığı bir modülü) hafızasından çağırarak, o çerçevede tanımlanan ilişkileri kullanması gerektiğini öne sürerek, dünyayı temsil etme sorununu, sistemin ortamında birçok çerçeve tanımlayarak çözmeye çalışmıştır. Her çerçeve tanımladığı ilişkilerle dış dünyayı yorumlayacağından, sorun hangi çerçevenin seçilmesi gerektiği sorununa indirgenir. Örneğin çerçeve, belirli bir oturma odası türünde bulunmak ya da

bir çocuğun doğum günü partisine gitmek benzeri durumları gösteren veri yapıları olabilir. Her çerçeveye belli bilgi türleri iliştirilmiştir. Bu bilgilerden bazıları çerçevenin nasıl kullanılacağı hakkında, bazıları çerçeveye özel beklentiler hakkındadır. (Minsky, 1974, s: 111-113)

Dreyfus, Heidegger'e göre insanın dünyada var olan bir varlık olduğunu vurgular. (Dreyfus, 1972, s: 168) İnsan, dünyaya fırlatıldığı ilk andan itibaren dünyayı deneyimleyerek anlamlandırdığından dünyayı sadece mantık kurallarıyla yorumlamaz. Dreyfus, insan davranışlarının türlü yaşantı ve deneyimlerle edinilmiş türlü bilgilerin sağduyuya dayalı bir birleşimi olduğu belirterek Minsky'nin çerçeve problemine sunduğu çözümün dünyadaki sonsuz değişik durum ve anlamı, sonlu çerçevelerde, sonlu bir kurallar dizisiyle tanımlamak olduğundan, önerdiği çözümünün başarılı olamayacağını savunmuştur. (Dreyfus, 2007, s: 248)

İnsan her zaman için dünyanın içindedir, dünya ile ilişki içinde dünyayı anlamlandırır. Dreyfus, öğrenmenin bir beceri ve alışkanlık kazanma süreci olduğunu belirterek, bu sürecin temsili ilişkilerden daha çok yönelimsel ilişki ağları kurma süreci olduğunu vurgular. Dünya ile ilk karşılaşmadan uzmanlığa doğru beş evrede kurulan ilişki ağları akıl yürütme süreçlerinden daha çok bakış açıları geliştirme becerileridir. (Dreyfus, Dreyfus, 1986, s: 16-36)

Dreyfus'un bir başka dikkat çektiği nokta ise Merlau Ponty'den yola çıkarak insanın bedeni ile dünyayla ilişkiye girişidir. Ponty, düşünme öncesi beden eylemine "motor yönelim" adını vermiştir; motor yönelim, yönelimselliğin temel formudur. Ponty'ye göre bedenin eylemi her zaman bir niyeti temsil etmez, beden nesnelerin çağrısıyla onlara doğru yönelir. Dreyfus bu durumda bedenin yöneliminin formel kurallar ve temsillerle açıklanmasının da imkansızlığına dikkat çekmiştir. (Dreyfus, 1972, s: 160-165)

Dreyfus'un beden vurgusu, yapay zeka çalışmalarının robotiğe kaymasında ve bedenlenmiş biliş kavramının doğmasında önemli bir rol oynamıştır. Yapay zeka araştırmacılarından Brooks, bedenin önemini vurgulayan "fiziksel temellendirme" hipotezini öne sürerek, dış dünyaya ait içsel temsillerin ancak bir robotun kendi temsillerini doğrudan kendi duyusal sensörleri aracılığıyla dış dünya ile ilişkiye girerek oluşturmasının mümkün olduğunu belirtmiştir. (Brooks, 1990, s: 6)

Dreyfus, güçlü yapay zekanın mümkün olmadığını, insan zihninin dünyayı anlamlandırmasının mantık kurallarıyla gerçekleşmediğini, insanların ancak dünyanın içinde var olarak dünyayı anlamlandırdığını, dolayısıyla belli kurallarla bir zihin

yaratılamayacağını düşünerek fenomenolojik öznenin zihinsel temsillerinin çeşitliliğine dikkat çekmiştir. İnsan zihninin yapısı bir şekilde biçimsel formüllere indirgense dahi insanın yaşam boyu kazandığı becerilerin sayısal verilere indirgenmesi imkansız olacaktır.

### 3.2.2 Çerçeve Problemi

Dennett, Dreyfus'un belirttiği problemin, açıkça tanımlanmamış olsa da "çerçeve problemi" olduğunu belirtir. (Dennett, 1984, s: 184) Çerçeve problemi ilk olarak McCarty ve Hayes tarafından sembolik sistemler göz önünde bulundurularak ortaya atılmış bir problemdir. Bir sistem çevre ile etkileşime girdiğinde sistemin içsel aksiyomlarından hangilerinin değişip hangilerinin sabit tutulması gerektiğinin belirlenmesini çerçeve problemi olarak tanımlanmıştır. (McCarthy, 1963) İçsel temsiline göre dünyayı ve dünyadaki ilişkileri tanımlayan bir program, dış dünyadaki durumsal değişikliklerde iç temsillerinde ve iç temsillerindeki ilişkilerde doğru bir güncelleme ile başa çıkmak durumundadır. Monotomik mantık aksiyomlarıyla tanımlanmış bir sistem söz konusu olduğunda sistemin içsel temsillerindeki herhangi bir değişiklik tüm aksiyomların değişmesiyle sonuçlanır. Her aksiyomun değişme sorunuyla başa çıkmak için tanımlanan çerçeve aksiyomlarıysa etkileşim sonucunda tüm aksiyomların değişmesini engellemek için sisteme getirilen ek aksiyomlar olarak tanımlanmıştır. Sembolik sistemlerde içsel temsillerin doğru güncellenmesini sağlayacak genel bir kural bulunması ise sağduyunun atalet sorunudur. Günümüzde sembolik sistemlerde monotomik olmayan mantık kullanılarak çerçeve sorunu kapalı sistemlerde belli ölçüde çözümlenmiştir. (Shanahan, 2009, s: 3-4) Ancak çerçeve problemi sembolik sistemlerin sorunlarıyla formüle edilmiş bir problem olmasına rağmen tüm yapay zeka mimarilerinin başa çıkmak zorunda olduğu bir problemdir. Çerçeve probleminin genel çözümünün mümkün olması ise güçlü yapay sistemlerinin mümkün olmasıyla aynı anlamı taşımaktadır.

Dolayısıyla Dennett çerçeve probleminin fenomenolojik bir problem olmadığını, çerçeve probleminin derin epistemolojik bir problem olduğunu savunur. Fenomenolojik yöntemin nesneleri bilincin nesneleridir. Ancak beynimizdeki hesaplamaların çok azı bilinçli içerikler oluşturmaktadır. Ayrıca fenomelojik yöntem çerçeve problemine birinci sahsın bakış açısından yaklaşmaktadır, bilim ise her zaman üçüncü sahsın bakış açısına sahiptir. Dennett'a göre yapay zekanın çerçeve

problemini çözmek ya da bilincin içeriklerini açıklayabilmek için fenomenolojinin ötesine geçmek zorundayızdır. (Dennett, 2011, s: 81) Fenomenolojiyi verili kabul eden Dennett, birinci tekil şahıs deneyimlerine ulaşmak için hetero-fenomenoloji yöntemini benimser. Bu yöntemde insanların bilincine ulaşmak için onların ifadelerine başvurmak zorunda olduğumuzu belirtir. Söz konusu olan denek davranışları tahmin edilebilecek, rasyonel bir taşıyıcı olarak kabul edilir. Bu deneğin düşüncelerine ulaşmaya çalışan kişi, deneğin betimlediği dünyayı değerlendirirken ne o dünyanın tamamen gerçek olduğunu kabul eder, ne de deneğin inançlarını yargılar, sadece betimlenen dünyayı inceler. Dennett'ın hetero-fenomenoloji yöntemi belli açılardan davranışsal bir yöntemdir. Ancak bu yöntemdeki davranışçılık deneğin tepkileri, düşünceleri, bakış açıları, duyguları ve iç gözlemleri de dahil olmak üzere beynin tüm içeriklerini kapsamaktadır. (Dennett, 2011, s: 90-96)

Dennett çerçeve problemini üç robot tasarımı üstünden betimler. R1 sadece kendi varlığını korumak için programlanmış bir robottur. R1'in enerjisi için gerekli olan pil ise bir odada üstünde aynı zamanda bomba olan bir vagondadır. R1 pilini kurtarmak için harekete geçer ve vagonu odadan dışarı çeker, bomba patlar. Problem R1'in vagonu dışarı çekerken bombayı da vagonla dışarıya çektiğini yani yaptığı hareketin odada tanımlanmış nesnelerin konumunu değiştirdiğini algılayamamıştır. Bir sonraki robot R1D1 yaptığı hareketlerin çevresindeki durumları nasıl değiştirdiğini hesaplayacak şekilde tasarlanır. R1D1 odaya girer, vagonu çektiği taktirde odada hangi koşulların değişip, hangi koşulların değişmediğini hesaplamaya başlar, örneğin vagon çekildiğinde duvarların rengi değişmeyecektir ve bomba patlar. Bu durumda mühendisler yeni tasarladıkları robota eyleminin çevresinde hangi durumlarla alakalı olduğunu sorgulayacak bir sistem yerleştirirler, yeni robot R2D2 vagonu çekmek üzere harekete geçmeden önce vagonu çekmesiyle alakasız yüzlerce olasılığı seri kurallar listesiyle yazılmış algoritmasıyla elerken bomba patlar. (Dennett, 1884, s: 181-182)

Özet olarak çerçeve problemi ilk olarak belirli bir durumda hangi bilgilerin o durumla alakalı, hangi bilgilerin durumla alakasız olduğunun tayin edilmesidir. Zeki bir ajanın bir durumda ancak o durumla ilgili bilgileri işleme koyması beklenir. Durumla ilgili bilgileri belirleyen bir ajanın hareketlerinin sonuçlarını tahmin etme yeteneği olması, hareketin sonucunda dış dünyanın temsillerinin iç temsillerini nasıl değiştireceği tahmin edebilmesi gerekir. Bu noktada Dennett insanların da hata yapabildiğini, bir ajanın beklenmedik bir sonuç karşısında şaşırması gerektiğini belirtir. Bir başka

sorunsa dünyanın robotun hareketlerinden de bağımsız olarak devamlı değişmesidir. Dolayısıyla robotun dış dünyaya ait içsel temsilleri dinamik olarak dünya ile birlikte değişebilmelidir. Aynı zamanda çerçeve problemi bir zeki ajanın neleri bilmesi gerektiğini ve bu bilgileri nasıl elde edeceğiyle de ilgilidir. Örneğin bir gece sandviçi hazırlarken nasıl sandviç hazırlayacağınızı, ekmeğe mayonez sürdüğünüz bıçağın katı bir nesne olduğunu ve birçok bilinçli olarak düşünmediğiniz bilgiyi kullanırsınız. Bilgilerimizin çoğu öğrenme sonucudur ancak doğuştan gelen bilgilerimiz de vardır, örneğin iki nesnenin aynı anda aynı yerde bulunamayacağını bilmemiz gibi. (Dennett, 1984, s: 187) Yapay sistemler, eğer ki belli bilgiler sisteme el ile konmazsa, bilgiyi edinmeye sıfırdan başlamak durumundadır. Dennett çerçeve problemini epistemolojik açıdan değerlendirerek çerçeve probleminin gerçek dünyada kapsamlı bir çözümünün güçlü yapay zeka olarak tanımlanan sistemlerin yaratılması için yeterli olduğunu öngörür.

Çerçeve problemi sadece yapay sistemlerin karşılaştığı bir sorun değildir. Herhangi bir hesaplamaya dayanan bilişsel modelde çerçeve problemine nasıl bir çözüm önerdiğini açıklamalıdır. Fodor, insan zihninin barındırdığı bilgi miktarından ötürü çerçeve probleminin hesaplanması mümkün olmayan bir problem olduğunu ve bilişsel bilimlerin araştırma programının ötesinde olduğunu iddia etmiştir. (Baars, Shananan, 2005, s: 159) Fodor'un zihni Turing makinesi gibi ele alan temsili zihin kuramında hesaplamalar semboller üstünden önermesel tutumlarla yapılmakta olduğundan sembollerin anlamının hesaplamaya bir katkısı bulunmamaktadır. Ancak bir durumda neyin durumla alakalı olduğu sembolün taşıdığı anlam ile yakından alakalıdır.

#### 3.2.2.1 Kapsamlı Çalışma Alanı Modeli ve Çerçeve Problemi

Kapsamlı çalışma alan modeli ise insan beyninin çerçeve problemini ne şekilde çözdüğüne bir açıklama getirdiğini öne sürer. İlk olarak insan beyni dış dünyayı iç temsillerinde betimler ve beyindeki iç temsiller dinamik olarak çevre ile her etkileşimde güncellenir. Kapsamlı çalışma alanı modelinin bir başka özelliği ise dış dünyadaki bir durumun beynin çeşitli modüllerinde paralel bir işlemeye tabi tutulmasıdır. Dolayısıyla dış dünyadan gelen bilgi, değer sistemlerinin kısıtlamaları altında hedeflere uygun bir şekilde çalışan dikkat mekanizmalarıyla değerlendirilmeye başlar. Bu şekilde her modül, kendi işlevine göre dış dünyanın durumuyla alakalı olan bilgileri işler ve her modül paralel olarak dış dünya ile alakalı olan bilgileri kapsamlı çalışma alanı

nöronları aracılığıyla tüm beyne iletir. Kapsamlı çalışma nöronları her modülün bilgilerini birleştirirken belli kombinasyonlar oluşturmaya başlar. Kapsamlı çalışma alanı nöronlarının çift yönlü iletişimi ile istatistiksel olarak öne çıkan aktivasyonlar yeniden alt modüllerde yankılanır. Bu şekilde hem dış dünyanın durumuna en alakalı bütün dinamik bir şekilde oluşturulmakla kalmaz, aynı zamanda sistem dinamik bir şekilde zamanda seri olarak dış dünyanın bir durumlarıyla etkileşime girer. Bir başka ifadeyle bilgiyi işleyen paralel modüller bilinçli olmayan hesapları yerine getirirken kapsamlı çalışma alanı nöronlarına ulaşarak tüm beyinde bütünleştirilen bilgi alt modüllerle döngüsel bir iletişim içinde olduğundan bilinçli düşüncenin seri yapısı ortaya çıkar. Seri yapıda işlenen bilgiden durumla en alakalı olanların bulunması paralel araştırma üstüne yapılan seri bir araştırma olarak düşünülebilir. Dolayısıyla kapsamlı çalışma alanının çerçeve problemine sunduğu çözüm bilinçle yakından alakalıdır ancak yapılan hesapların çoğu paralel modülerde bilince çıkmaksızın yapılır. (Baars, Shananan, 2005, s: 157-174)

Eğer R1D2 robotunun işlemcisi dış dünyanın durumlarını iç temsillerinde dinamik olarak oluşturan birçok modülden meydana gelen bir işlemci olsaydı, R1D2 odaya girdiğinde ilk olarak hedefi doğrultusunda pil ile alakalı modüller odadaki pil ile alakalı bilgileri ön plana çıkaracaktı. Ama aynı zamanda R1D2'nin duyusal algı modülleri iç temsillerinde pilin yanında duran bir bomba beklemediği için bombanın varlığı da R1D2'nin sisteminde yankılanacak ve sistem bomba bilgisini yeniden alt modüllere göndererek tüm modüllerde bomba ile alakalı tecrübeler arasında önemli olanların derlemesinin yapılmasını sağlayacaktı. Bu şekilde R1D2 hesap kapasitesi ve daha önceki bilgi birikimine göre seri bir çözüm üretilmeye çalışılacaktı. Yaratılacak olan çözümün işe yaraması ise R1D2'nin içsel modellerinde temsil ettiği bilgilere bağlı olacaktı. Bu noktada ise insanların doğuştan sahip olduğu veya gelişimlerin erken döneminde edindikleri, örneğin birbirine yapışık iki cismin birlikte hareket ettiği bilgisi gibi bilgilere R1D2 sahip olup, olmadığı önem kazanacaktı.

Kapsamlı çalışma alanı modelinin çerçeve problemine sunduğu çözümde iki önemli nokta beynin dış dünyanın temsillerini dinamik olarak oluşturmasının yanı sıra bilginin ilk olarak değişik modüllerde paralel olarak işlenmesidir. Tekrarlayan sinir ağlarının söz konusu bilginin dinamik temsillerini oluşturması bu ağlar çerçeve problemine bir çözüm sunma ihtimalleri açısından önemli bir avantaj sağlamaktadır. Dennett özellikle planlanmış durumlarda insanların hata yapabildiğini ve şaşırabildiğini belirtmiştir. (Dennett, 1984, s: 193) Kapsamlı çalışma alanı modelinin en temelde dış duyusal algı

ile içsel temsilleri arasında istatistiki bir düzeltme yapıyor olması ve istatistiki derleme sonucu oluşan seri döngüsel yapı aynı zamanda bir modülün hatasının, söz konusu durumla yanlış ilişki kurmasının nasıl genel bir hataya neden olduğunu açıklamaktadır. Hataların sistemde yankılanması ise şaşkınlık olarak tanımlanır. Özellikle zamansal fark öğrenmesi mekanizmasında hata ne derece büyük ise şaşkınlık o derece büyük olur ve sistem hatasından öğrenerek iç temsillerini yeniden yapılandırır.

Kapsamlı çalışma alanı modeli dinamik olarak oluşturulan iç temsiller aracılığıyla beynin belli bir anda çerçeve problemini nasıl çözebileceğine dair bir fikir sunmaktadır. Ancak iç temsillerin ilk anda hangi zorunluluklarla oluşturulduğu sorusunu yanıtlamamaktadır. İnsan beyni en baştan dünyanın tüm özelliklerinin temsillerini oluşturmamakta dünyanın belli özelliklerine dair temsiller oluşturmaktadır. Biyolojik bilişsel sistemler ise her zaman için çevre ile etkileşim içinde ortaya çıkmış sistemlerdir. Biyolojik bir biliş dinamik olarak oluşturmuş ve oluşturulmakta olan iç temsiller üstünden yapılan hesaplamalar olarak kabul edilse bile bilişsel sistemin yapısının ortaya çıkışında çevre ve organizmanın etkileşimleri belirleyicidir. Damasio, çevrenin hangi temsillerinin organizma için önemli olduğunu belirleyen mekanizmaların ise evrimle ortaya çıkmış duygu mekanizmaları ve duygu mekanizmalarından türemiş olan hisler olduğunu belirtmiştir. (Damasio, 2020, s: 27)

Edelman insan beyninin evrimsel olarak üç aşamayla, gelişimsel seçilim, deneyimsel seçilim ve seçilmiş nöronal olayların uzay-zaman bağlantısını temin etmek için dağıtık nöron grupları arasındaki ve içindeki karşılıklı bağlantılar üzerinden sinyalleşen yeniden girişli süreç ile şekillendiğini belirtir. (Edelman, 1989, s: 11) Gelişimsel seçilim beyin anatomisinin ilk şekillenişinin gerçekleştiği embriyo döneminde gerçekleşir. Bu dönem ilk olarak genetik kalıtımla sınırlanmıştır ancak sinaps düzeyindeki bağlantılar büyük oranda somatik seçilim yoluyla kurulur. Gelişme esnasında nöronlar dallanarak sayısız çıkıntılarını birçok yöne dağıtırlar, bu dallanma bağlantı örüntülerinde geniş çaplı bir çeşitlilik üretir ve bu çeşitlilik somatik seçilimle ayıklanır. Sinir sisteminin son yapısal hali sağlandığında üretilmiş olan nöronların hatırı sayılır bir yüzdesinin öldüğü bilinmektedir. (Edelman, Tonini, 2019, s: 125-126) Deneyimsel seçilim, davranışsal deneyimin bir sonucudur. Kortekste bazı bağlantılar güçlenirken, bazı bağlantılar güçsüzleşir, hatta yok olurlar. Bu seçilim süreci, organizmanın hayatta kalmasını sağlayarak evrimin ilk adımlarında oluşmuş değer sistemleri tarafından sınırlandırılmıştır. (Edelman, Tonini, 2019, s: 126) Yeniden giriş ise beyindeki çeşitli

örüntü ve haritaların baştan etiketlenmemiş bir dünyanın nesne ve olaylara ayrılmasını ve korteksin dünyayı anlamlandırma sürecini içerir. Yeniden giriş çeşitli duyusal ve motor olayların uzaysal zamansal eşgüdümünün gerçekleştiği merkezi mekanizmadır. (Edelman, Tonini, 2019, s: 126) Dolayısıyla insan beyni dış dünya durumlarını iç temsillerinde oluştururken hem organizmanın çevreyle dinamik etkileşimini kullanmakta hem de evrim sürecinde gelişmiş mekanizmalardan yararlanmaktadır. Bir başka ifadeyle yapay sistemler çevreyle etkileşmeye bir modelden yoksun bir şekilde başlarken, insan beyni doğuştan belli niteliklere sahip olarak çevreyle etkileşmeye başlar**.** Doğa kendi bilişsel tekerleklerini evrim sürecinde kendisi yaratmıştır. (Dennett, 1984, s: 201)

#### 3.2.2.2 Yapay Sinir Ağları ve Çerçeve Problemi

Günümüzde tekrarlayan sinir ağları içinde bulundukları dünyanın durumlarını dinamik olarak iç temsillerinde oluşturmakta, derin pekiştirmeli öğrenme ise çevreyle etkileşim içindeki ajanın belirli hedeflere yönelik hareket ederken çevresini araştırmasını ve hedefiyle alakalı durumları öğrenmesini sağlamaktadır. Ancak günümüzde insan seviyesinde yapay sistemler yaratılmasının önünde ciddi problemler vardır. Bu problemler en temelde ajanın neleri bilmesi gerektiğini ve bu bilgileri nasıl elde edeceğiyle ilgilidir. İnsan zihninin sahip olduğu nedensellik, genelleme ve dolayısıyla soyut kavramlar oluşturabilme yetileri bağlantıcılık temelli modellerin sahip olmadığı kavram ve yetilerdir.

Bağlantıcılık temelli modeller bilişsel sistemlerin değişik düzeyleri olduğunu ve nöronal hesaplamaların alt-sembolik seviyede olduğu teziyle sistemler üretmektedir. Günümüzdeki derin ağların ancak istatistiki bilgiler ile sınırlandırılmış genelleme yeteneğine sahip olması bu sistemleri insan zihnini sahip olduğu birçok özellikten mahrum bırakır. Ancak yapay sinir ağı araştırmacıları bağlantıcılık çerçevesi içinde üst düzey mimariler ile söz konusu sorunların teorikte çözülebileceği inancıyla araştırmalarına devam etmektedir. Günümüzde derin ağlar ve derin pekiştirmeli öğrenmeyi birleştiren sistemler kapalı çevrelerde hatırı sayılır başarılar elde etmişlerdir. Bir sistemin kapalı olması ajanın etkileştiği çevrenin durumlarının sonlu bir durum alanı kümesiyle tanımlı olması anlamına gelmektedir. Açık sistemlerde ajanın daha önce hiç karşılaşmadığı çevre durumlarıyla karşılaşması olasıdır. Dolayısıyla açık bir sistemde ajanın çevre durumlarıyla baş etmesi için gerekli olan

yeteneklerden biri çok az sayıda örnekten hatta tek bir örnekten öğrenme yetisidir. Yapay sistemlerin öğrenmek için belli bir istatistiki derleme yapma zorunluluklarının nedeni bu sistemlerin genel olarak tek bir hedefe yönelik hareket edecek şekilde tasarlanmasıdır. İnsanlar yeni bir durumla karşılaştıklarında hali hazırda dünya hakkında çok fazla şey bilmekte ve yeni bir durumla karşılaştıklarında sahip oldukları bilgiyi yeni duruma göre sistematik bir şekilde genelleyebilmektelerdir. Geniş bir genelleme yeteneği sadece az sayıda örnekten öğrenmekle alakalı değildir, çevrede oluşan genel bir değişikliğe adapte olmak için de gereklidir. Örneğin odanın ışığı kapatıldığında insan zihni bu durumla başa çıkabilirken derin bir ağ çevrede değişen tek parametrenin ışık seviyesi olduğunu anlamayacak, aydınlıkta öğrendiği ilişkileri karanlık ortama genelleyemeyecektir.

İnsan zihni henüz dil gelişiminden önce gözetimsiz şekilde fizik yasalarını öğrenmeye başlar. Henüz üç aylık bebekler nesnelerin yapısını algılar, nesneler hareket ederken süreklilik, yapışkanlık ve temas yasalarına uymasını beklerler. (Pinker, 2016, s: 368) Bebeklerin erken dönemde edindikleri bu bilgiler sembolik sistemlerde tanımlanması mümkün olmayan, derin ağların ise henüz temsilini oluşturamadığı bilgilerdir. Dennett, dünyanın işleyişine dair oluşturduğumuz sezgisel bilgilerin bilinçli düşünmeksizin bilgimiz bilgiler olduğunu, örneğin iki cismin aynı anda aynı yerde olamayacağı bilgisini bilinçli olarak düşünmeye gerek duymaksızın bildiğimizi vurgulamıştır. (Dennett, 1984, s: 188) Aynı şekilde nedensellik kavramı da insan zihninde erken dönemde oluşan soyut bir kavram ve yapay sistemlerin sahip olmadığı bir kavramdır. Nedensellik dünyaya dair verileri istatistiki dağılımlarının genellemesi ile oluşturulan bir kavram olmaktan daha çok ajanın dünya ile etkileşimiyle ortaya çıkan bir kavramdır. Nedensellik aynı zamanda nesneleri tanımlamamızla da alakalıdır. Bir sandalyeyi tanımlayan sandalyenin görüntüsünden daha çok sandalyenin oturulma özelliğidir. Bir sandalyeye sandalye denmesinin nedeni sandalyeye oturulmasıdır. Çok az sayıda örnekten öğrenmenin ötesinde insan zihni belirli durumlarda belli hareketlerin sonucunu hayal etme yetisine de sahiptir ki bu yapay sistemlerin sahip olmadığı başka bir özelliktir.

Bilişsel bilimci Kahneman insan zihninin iki sistemden oluştuğunu öne sürmüştür. (Kahneman, 2011) Sistem 1 hızlı ve sezgiseldir, sistem 1'in günümüzdeki derin öğrenmenin oldukça başarılı olduğu dış dünyanın istatistiki derlemelerini oluşturan sistem olduğu varsayılabilir. Sistem 2 ise sistem 1'in üstüne kurulmuş seri yapıdır, bir başka ifadeyle rasyonel düşünen, planlama yapabilen, düşündüklerini dil ile ifade

edilebilen sistemdir. Sembolik sistemler sistem 2'yi sistem 1 olmaksızın taklit etmeye çalışmıştır. Fodor'un temsili zihin kuramı da sistem 1 ile sistem 2'nin etkileşimini irdelememiştir. Günümüzde derin öğrenme sistem 1 ve sistem 2 arasındaki boşluğu doldurmaya çalışmaktadır. Kapsamlı çalışma alanı modeli sistem 2'nin oluşmasında ilk aşamada sistem 1'e ait bilginin önemli kısmının seçilmesini ve bu bilginin seri şekilde işlenmesini açıklamaktadır. Derin pekiştirmeli öğrenme tekrarlayan sinir ağlarıyla birleştirildiğinde sistemin çalışan hafızasındaki aktivasyon desenini öğrenerek çok basitte olsa düzenlilikleri öğrenebiliyor olması veya pekiştirmeli öğrenme sistemlerine epizodik hafıza birimi gibi işlev görecek basit eklemeler yapılmasıyla sistemin ilk defa karşılaştığı bir duruma geçmiş deneyimlerinden bir çözüm bulmasının sağlanması, insanlarda bulunan birçok hafıza mekanizmasının yapay sistemlerin eksiklerini tamamlamak için elzem olduğu düşüncesini akla getirmektedir. Günümüzde epizodik belleği biyolojik olarak makul olmayan teknikler ile taklit edilmesi ve aynı şekilde sistemlere dikkat mekanizmalarının eklenmesi dahi yapay sinir ağları mimarilerinin beklenenin ötesinde başarı elde etmesini sağlamıştır. Ayrıca yapay zekada kredi tahsisi olarak tanımlanan, önceden olmuş bir olaydan sonradan çıkarım yapmak gibi yetiler de hafıza ile ilişkili problemlerdir. Epizodik bellek veya dikkat ise bilinçle iç içe geçmiş yapılardır.

Dennett bilginin ifade edilmesinin ve kullanılabilir hale getirilmesinin dil ile gerçekleştiğini ifade eder. (Dennett, 1984, s: 189) Dilin ortaya çıkışı ise kültürün ortaya çıkışıyla paralel bir süreçtir. Dolayısıyla insanların dünyasında iş görecek yapay bir sistemin çerçeve problemine getireceği çözüm insan kültürünü de hesaba katmak zorunda kalacaktır.

Dennett çerçeve problemine epistemolojik bir problem olarak yaklaşmış ve çerçeve probleminin fenomenolojik bir problem olmadığını vurgulamıştır. Çerçeve problemine fenomemolojik açıdan yaklaşmak öncelikle beynin birçok hesabını dışarıda bıraktığı için hatalı bir yöntemdir. Fenomenoloji sadece zihnin bilinçli içerikleriyle ilgilenir, ayrıca bilinç kavramı yönelimsel duruşu benimsediğimizde önümüze çıkan bir kavramdır. Dennett, yapay bir sistemin bilinçli olmasından veya bilinçle yakından alakalı yetileri olmasından daha çok sistemin çerçeve problemini çözecek yetileri olup, olmadığıyla ilgilenmekte ve ontolojik tartışmaları çerçeve probleminin dışında tutmaktadır. Dennett'ın tanımladığı çerçeve problemini çözen bir ajan ise insan kültürü içinde fenomenolojik bir özneye tekabül edecektir ancak kullanılan yöntem heteroofenomenolojik yöntem olduğundan bu öznenin fenomenolojik deneyimlerinin

kendisine sunulduğu bir insanınki gibi qualia uzayı olup olmadığını ancak sistemin rapor ettikleriyle ölçme ve değerlendirme imkanımız olacaktır.

## 3.3 BİLİNCİN ZOR PROBLEMİ

Dennett bilinci epistemolojik bir problemin çözümünde işlevsel bir rolü olan bir kavram olarak kullanmakta ve bilincin özneye sunulan temsillerinin bir başka ifade ile öznenin *qualia'*sının özneye hangi koşul ve şekillerde sunulduğunu söz konusu epistemolojik problemin dışında bırakmaktadır. Ancak Dennett'ın aksine nörobilim ve yapay zeka araştırmacılarının büyük bir kısmı çerçeve probleminin çözümü için önce bilincin doğasının anlaşılmasının gerekli olduğunu düşünmektedir. Nörobilim araştırmaları bilincin ne olduğunu ve bilinçli mekanizmaların tam olarak nasıl işlediğini çözmemiş olsa da nörobilim alanındaki araştırmacılar, bilinç probleminin beynin anatomik ve fonksiyonel yapıları anlaşıldıkça çözüme kavuşacağına inanmaktadır. En nihayetinde nörobilim bilincin temelinde beynin belli özellikleri olduğunu veya zihnin beynin karmaşıklığı sonucu zayıf beliren bir özellik olduğunu varsaymaktadır.

### 3.3.1 Zor Problem

David Chalmers ise bilimin ancak bilincin belli özelliklerini kavrayabileceğini iddia ederek bilinç problemini iki kısma ayırmıştır. Chalmers'a göre nörobilim araştırmalarının çözmeye çalıştığı sorunlar, bilincin kolay problemleridir. Kolay problemler kolay değildirler ama belli bir yöntemle çözülebilecek problemlerdir. -Chalmers'a göre duyusal uyaranların ayırt edilmesi, beynin bilgiyi nasıl entegre ettiği, uzun süreli bellek oluşumu, düşünme ve karar verme mekanizmaları, zihinsel durumları bildirme, dikkat mekanizmaları işlevsel bir bakış açısıyla belli fiziksel bağlantılar ve mekanizmalarla açıklanabilecek problemlerdir.- Chalmers bilincin zor problemini neden öznel deneyimlerimiz olduğu, öznel bilinç içeriklerimizin fiziksel bir sistem tarafından nasıl yaratıldığı sorularıyla tanımlar. (Chalmers, 1995, s: 202) Zor problemin kolay problemleri çözen metotlarla cevaplanamayacağını düşünür.

Uyanık olduğumuz veya rüya gördüğümüz her anda bize eşlik eden bilinç deneyimi zihinsel temsillerimizi içermektedir. Bu temsiller dışarıdan ulaşılamazlar ve bize özeldirler. Ekşi elmanın tadı, turkuaz bir denizi seyretmek, diş ağrısı gibi zihinsel durumların tamamı deneyimleyen kişiye özgü durumlardır. Herhangi bir deneyimin

kişiye özel yanına deneyimin fenomenal yönü denmektedir. Fenomenal deneyimi oluşturan zihinsel temsiller ise deneyimleyenin *qualiasını* oluşturur. Bir deneyimin *qualiası*, deneyimi deneyimleyen için deneyimleyen olmanın nasıl bir şey olduğunu da ifade eder. *Qualia* duyusal algılar (renkler, sesler, kokular, sıcak, soğuk gibi duyumlar), deneyimleyenin vücuduna ait hisleri (açlık, tokluk, susuzluk, ağrı) ve deneyimlenen duyguları (sevgi, nefret, öfke, korku) içerir. (Dennett, 2011, s: 55)

Chalmers, fenomenal deneyimlerimizin, sahip olduğumuz *qualia* uzayının hayatımıza anlam katan bir özellik olduğunu belirtir. Ancak fenomenal deneyimlerimizin beynimizin fiziksel özellikleriyle bağlantılı olsa bile beynin fiziksel özelliklerinden yola çıkarak anlaşılamayacağını savunur. Fenomenal deneyimlerin beynin fiziksel süreçleriyle alakalı olduğu benimsenmiş bir düşüncedir, beyin süreçlerimiz fenomenal deneyimlerimizi yaratır. Zor problemin çözümü fiziksel süreçler ve bilinç arasındaki ilişkiyi, öznel deneyimin bu süreçlere neden ve nasıl eşlik ettiğinin açıklamasını içermelidir. (Chalmers, 1995, s: 202)

Chalmers, fenomenal deneyimleri açıklamaya çalışan tüm girişimlerin sınıflandırmasını yapmıştır. Chalmers'ın sınıflandırmasında ilk ayrım indirgemeci ve indirgemeci olmayan yaklaşımlar arasındadır. Genel olarak indirgemeci yaklaşımlar, zihnin beynin fiziksel yapısının bir sonucu olduğunu savunurken, indirgemeci olmayan yaklaşımlar zihnin fiziksel dünyadan ayrı ontolojik bir yapı olduğunu savunmaktadırlar. Chalmers, ilk olarak beyin zihin özdeşliğine karşı olan üç argümanı -açıklayıcı boşluk, tasarlanabilirlik ve bilgi argümanlarını- inceler. İndirgemeci yaklaşımlar, indirgemeci olmayan yaklaşımları destekleyen bu argümanların hepsine bir açıklama getirmek zorundadır. (Chalmers, 2003, s: 4)

Açıklayıcı boşluk argümanı: İndirgemeci yaklaşımlara karşı çıkan argümanlardan ilki, açıklayıcı boşluk argümandır. 1983 yılında Levine tarafından öne sürülen bu argümanın temelini davranışsal ve bilişsel fonksiyonlar gibi fiziksel süreçlerle açıklanabilen bilincin kolay problemlerini anlamanın bilincin yapısını anlamak ve zor problemi çözmek için yeterli olmadığı düşüncesi oluşturur. (Levine, 1983, s: 360) Levine, Kripke'nin materyalizme karşı öne sürdüğü kartezyen argümanın iki ana iddiasından yola çıkarak, Kripke'nin metafiziksel argümanlarını epistemolojik argümanlara çevirmeyi hedeflemiştir. Levine, bu yöntemle *qualia*nın fiziksel yaklaşımlarla açıklandığı taktirde bu açıklamaların fenomenal deneyimin doğasını tanımlamayacağını göstermiştir. Fiziksel açıklamanın fenomenal deneyimi tam açıklayamıyor olmasını ise "açıklayıcı boşluk" olarak tanımlamıştır.

Kripke’nin iki ana iddiasından ilki katı tasarıma (bir niteliğin tüm olası dünyalarda aynı şeyi göstermesi) sahip tüm özdeşlik önermelerinin çift yönlü doğru olmasının ancak ve ancak bu önermelerin tüm olası dünyalarda doğru olmasıyla mümkün olduğudur. İkinci iddiası ise fizyolojik önermelerin yanlış olduğunu düşünmenin mümkün olduğudur. Levine, fizyolojik özdeşlik önermelerinin yanlış olduğunu düşünmenin mümkün olduğunu belirlemek için bir yöntemin olmadığını ancak bu önermenin fenomenal deneyimleri açıklamak adına ciddi bir epistemolojik boşluk yarattığını belirtir. (Levine, 1983, s: 355)

Levine, iki fiziksel özdeşlik argümanı tanımlar : 1- Acı X nöronunun ateşlenmesidir. 2- Isı moleküllerin hareketidir. Isının molekül hareketlerinden bağımsız tanımlanabileceği olası dünyalar yoktur. Ancak ısının hissi olan sıcaklığın aynı olmadığı olası dünyalar vardır. Levine sıcaklık ve ısının farklı olduğunu ve ısının her zaman moleküllerin hareketi olup, sıcaklığın moleküllerin hareketleri aracılığıyla tanımlanabileceğini ancak aynı mantığın acıya uyarlanamayacağını çünkü hem acı hissinin hem de acı deneyiminin acı olarak tanımlandığına dikkat çeker. Zihinsel durumlar, beynin fiziksel durumlarından daha çok beynin işlevsel durumları olarak tanımlandığında da fenomenal durumların açıklamasına bir katkı yapmaz. Acı bir f durumunda olmak olarak tanımlanırsa, f durumunun acıya özdeş olmadığı bir dünyanın mümkün olup olmadığı sorgulanmalıdır. (Levine, 1983, s: 355) Levine, bu noktada Block’un Çin deneyini öne sürer. Block, her Çinlinin bir nöronu temsil ettiği bir düzenekte, beynin acı esnasındaki aktivasyon durumunu Çinlilere taklit ettirdiğimizde Çin’in ülke olarak acı çekiyor olduğunu düşünmenin mantıksız olduğunu belirtmiştir. (Levine, 1983, s: 356) Levine, X nöronunun ateşlenmesiyle ya da acıyı oluşturan f durumunda acıdan sakınmak için gerekli olan işlevsel tepkileri görebileceğimizi, X nöronun ateşlenmesi ve beynin f durumunda olması ile acı deneyimi arasında bir korelasyonda olabileceğini ancak acı deneyimiyle bağlantılı X nöronunun ateşlenmesinin veya f durumunda olmanın fenomenal acı deneyiminin nasıl bir deneyim olduğunun bilgisini içermediğini belirtmiştir. Dolayısıyla acının fiziksel açıklaması ile acı hissinin fenomenal deneyimi arasında epistemolojik bir açıklama boşluğu vardır. Bir acı deneyimin tüm nöronal korelasyonlarını bilsek bile bu korelasyonlar acı deneyimi hissinin nasıl bir his olduğunu açıklayamayacaktır. (Levine, 1983, s: 360)

Chalmers, açıklayıcı argümandan yola çıkarak, fiziksel olanın sistemin yapısını ve işlevini açıkladığını, sistemin yapısını ve işlevini açıklamanın bilinci açıklamak için

yeterli olmadığını, dolayısıyla bilincin fiziksel bir açıklaması olmadığını belirtir. (Chalmers, 2003, s: 5)

Tasarlanabilirlik argümanı : Tasarlanabilirlik argümanının ilk çıkış noktası, Descartes'ın hayvanların bir otomat olduğunu, insan davranışlarının da aynı şekilde fiziksel olarak açıklanabildiğini öne sürmüş olmasıdır. On dokuzuncu yüzyılda bilimsel düşüncenin ağırlık kazanmasıyla, her fiziksel tepkinin fiziksel bir nedeni olduğu ve fiziksel dünyanın nedensellik ilkelerine uyduğu görüşü ağırlık kazanmıştır. Fiziksel dünyada nedensellik ilişkileri geçerliyse bilincin bir işlevi olmadığı, hatta bilincin beynin işlevinin bir yan ürünü olduğu düşüncesi Huxley tarafından ortaya atılmıştır. Huxley, tüm fiziksel olanın, insan davranışları da dahil olmak üzere nedensellik yasalarına uyduğunu, bilincin etkisiz bir yan ürün olduğunu savunmuştur. (Dennett, 2011, s: 468) Huxley'in düşüncesinin bir başka yorumu eğer bilincin bir işlevi yoksa fiziksel olarak bilinçli insanlarla aynı davranışları sergileyen ama bilinçten yoksun zombilerin varlığının mümkün olduğudur.

Chalmers, tasarlanabilirlik argümanını kendi önermeleriyle ifade ederek, zombilerin düşünülebilir olduğunu, zombiler düşünülebilir ise zombilerin varlığının metafiziksel olarak mümkün olduğunu, eğer zombiler metafiziksel olarak mümkün ise bilincin fiziksel olmadığını belirtir. (Chalmers, 2003, s: 6)

Bilgi argümanı : Bilgi argümanı bir insan herhangi bir fenomenal bir deneyimle ilgili tüm fiziksel durumları bilse bile, eğer söz konusu fenomenal durumları deneyimlememiş ise bu fenomenal deneyimle ilgili bazı durumların bilgisine ulaşamayacağını savunur.

1982 yılında Jackson, fiziksel durumların bilgisinin fenomenal durumların deneyimlenerek elde edilen bilgisini kapsamadığını gösteren bir düşünce deneyi tasarlamıştır. Bu deneyde Mary çok başarılı bir bilim insanıdır, ancak siyah beyaz bir odadaki siyah beyaz bir televizyondan dünyayı incelemeye mahkum edilmiştir. Mary görmenin fizyolojisi üstüne uzmanlaşmıştır ve renklerle ilgili tüm fiziksel bilgiyi bilmektedir. Mary siyah beyaz odasından çıkarılıp, ilk kez renkleri gördüğünde, renklerin fenomenolojik bilgisini elde ettiğinde kuşkusuz renklerle ilgili yeni bir bilgi edinmektedir. Jackson, bu durumda fiziksel bilginin tüm bilgiyi kapsamadığını, fiziksel bilgi ile fenomenal deneyimin bilgisi arasında fark olduğunu, fenomenal bilginin, fiziksel bilgiden türetilemeyeceğini belirtir. (Jackson, 1982, s: 291-294)

Chalmers, bilgi argümanını önermelerle ifade eder; Mary, renkler hakkında tüm fiziksel bilgiye sahiptir. Mary renklerle ilgili tüm bilgiye sahip değildir dolaysıyla fiziksel bilgi tüm gerçekliği kapsamaz. Bu önermelerin diğer bir ifadesi ise fenomenal hakikatlerin fiziksel hakikatlerden türetilemediği dolayısıyla bilincin fiziksel hakikatlerden farklı bir olgu olduğudur. (Chalmers, 2003, s: 7)

### 3.3.2 Bilincin Zor Problemine Yaklaşımlar

Chalmers, açıklayıcı boşluk, tasarlanabilirlik ve bilgi argümanlarının ortak noktasının fiziksel olanla fenomenal olan arasındaki bilgi boşluğunu vurgulamaları olduğunu belirtir ve indirgemeci yaklaşımların bu üç argümanı yadsıdıklarını iddia eder. Chalmers, indirgemeci yaklaşımları bu argümanlara yaklaşımlarındaki farka göre üç alt gruba ayırır. (Chalmers, 2003, s: 9)

Chalmers'a göre A tipi indirgemeci yaklaşımlar fiziksel ve fenomenal doğrular arasında boşluk olmadığını, zombilerin mümkün olmadığını ve Mary'nin kırmızı hakkındaki fenomenolojik bilgisinin eksik olduğunu kabul eder; bilincin kolay problemlerinin çözülmesiyle bilincin zor probleminin kendiliğinden bir çözüme kavuşacağını savunurlar. Chalmers, A tipi materyalistlerin bilincin işlevinin ötesinde bir açıklama getirilecek derinliğe sahip olmadığını düşündüklerini belirtir. (Chalmers, 2003, s: 9) Elemeci materyalizm ve indirgemeci materyalizm Chalmers'ın sınıflandırmasına göre A tipi indirgemeci yaklaşımlardır.

Elemeci materyalizmi savunanlar, bilim tarihinde ampirik olarak araştırılmaya başlanan birçok kavramın, araştırmanın sonunda ya kullanımdan kalktığını ya da anlamının değiştiğini vurgularlar. Çoğu kavram günlük konuşma diliyle şekillenmiştir. Bilim bu kavramlar üstüne düşünmeye başladığında, bu kavramların düşünülmüş olan anlamından çok farklı anlamları olduğu ortaya çıkar. Örneğin atom kavramı bölünemez en temel parçacık anlamına gelirken bilim atomun daha temel parçacıklardan oluştuğunu, atomların parçalanabilir olduklarını göstermiştir. Atom kelimesi hala kullanılmakla beraber taşıdığı anlam radikal biçimde değişmiştir. Elemeci materyalistler, günlük düşünmenin bilimsel düşünceden ayrıldığının, günlük hayatımızda kullandığımız sağduyuya bağlı halk kuramlarının genelde bilimsel kuramlar ile çeliştiğinin de altını çizerler. Halk kuramlarına bir örnek vermek gerekirse McCloskey'in araştırması günümüzde dahi insanların hareket üstüne düşünürken Newton mekaniğindense sağduyu ile temellenmiş Aristoteles fiziğini göstermiştir. Bir

deneyde üniversite öğrencilerinden oluşan deneklerden koşarak bir hedefi golf topuyla vurmaları istenmiştir. Denekler topun Aristoteles fiziğinde olduğu gibi düz düşeceğini düşünerek topu hedeflerinin üstüne gelmeden topu bırakmamışlardır. Bu deney ve bu tarz deneyler, bilim toplulukları dışındaki insanların sezgilerine dayalı fizik kurallarını tercih ettiklerini, problemler üstüne düşünürken halk kuramları ile düşündüklerini göstermiştir. (Churchland, 2019, s: 321) Halk kuramları, bilinçli bir kuramsallaştırma sürecinin ürünleri değildir ancak kişilerin inançları doğrultusunda iş görürler. Churchland halk kuramlarıyla araştırmaların ön bilimsel kuramlarının arasında bir sınır çizmenin zor olduğunu ancak bilimsel kuramlar geliştikçe halk kuramlarına ait kavramlarının çok farklı bir bakış açısıyla ele alınmaya başlandığını belirtir. (Churchland, 2019, s: 312)

Aynı şekilde elemeci materyalizm inanç, arzu, irade kavramları gibi bilinç kavramının da halk psikolojisinin bir kavramı olduğunu ve bu kavramların halk psikolojisinde taşıdıkları anlamların bilimsel bir zihin teorisinde değişeceğini, hatta zihne dair kullanılan bazı kavramların eleneceğini belirtir. Halk psikolojisi, insan davranışlarını açıklar ve öngörüde bulunurken günlük kavramları genellemeye meyleder. Bu genellemeleri ise insanların kendilerini nasıl kavramsallaştırdığı belirler. Churchland felsefecilerin analiz ettiği inanç ve istek kavramlarının halk psikolojisinin temel kavramları olduğunu öne sürer. (Churchland, 2019, s: 332) Churchland, özellikle zihin felsefesinde tercih edilen araştırma yönteminin zihinsel olandan bahsedilirken kullanılan yaygın kavramları analiz etme yoluyla cevaplara ulaşmak olduğunu belirtir. Bu analiz yöntemi neyin düşünülebilir ve neyin hayal edilebilir olduğunu sorarak, bir şeyin soru konusu olan halihazırdaki anlamlarla bağdaşıp bağdaşmadığını tahlil edilerek ve eldeki sorunda sıradan insanların eldeki sorunla ilgili kelimeleri ya da felsefi kavramları nasıl kullandığını inceleyerek gerçekleştirir. (Churchland, 2019, s: 300) Örnek vermek gerekirse zihnin en temel kavramlarının inanç ve istekler olduğunu düşünen felsefeciler, inanç kavramını analiz ederken, inancın belli semboller ve mantık kurallarıyla alakalı olduğunu belirtmiş ve inanç tanımını bu gözlemleri doğrultusunda yapmıştırlar. Bu şekilde inançların bilişsel bir sistem oluşturarak belli kurallara uyduğunu ve inançların semantik olarak yorumlanabileceğini öne sürmüştürler. Fodor'un "düşünce dili teorisi", Churchland'a göre halk psikolojisi kavramlarının analizine dayanmaktadır. Churchland, halk psikolojisinin kavramlarının analizinden beynin yapısı hakkında bir sonuca ulaşmak yerine beynin yapısından belli kavramlara ulaşılması gerektiğini vurgular. Nörobilime

göre inanç, öğrenme sürecince oluşan belli bir nöron ağındaki aktivasyonların ağırlıklarınca belirlenen bir temsildir, dil ise bu temsilin ifadesidir.

Elemeci materyalistler, bilinç kavramının halk psikolojisinin bir kavramı olduğunu, bilimsel anlamda bilincin tanımının, bilinç kavramının ifade ettiklerini analiz ederek yapılamayacağını, bu tip analizlerin ancak yanlış yönlendirmelere yol açacağını belirterek bilinç kavramının bilimsel bir araştırma programından elenmesi gerektiğini savunurlar. Beyin durumlarının belli zihinsel durumları açıklamıyor gibi gözükmesinin nedenini kullandığımız halk psikolojisi kavramlarının nörobilim için uygun olmamasına bağlarlar. Elemeci materyalistlere göre bilim, günlük konuşma dili kavramlarını ya tamamen eleyerek ya da anlamlarını değiştirerek bir zihin teorisine ulaşacaktır.

İndirgemeci materyalistler ise bilincin ve zihinsel temsillerin gerçek olduğunu, ancak zihinsel durumların beynin bir durumuna tekabül ettiğini iddia eder. Kısacası indirgemeci materyalistlere göre beyin ve zihin özdeştir; öznel psikolojik gerçeklik nesnel nöronal bir gerçekliğin bir tezahürüdür; bilinç bir beyin sürecidir. İndirgemeci materyalistler, ileride nörobilimin zihinsel temsiller ile nöronal temsiller arasında bir köprü oluşturacağını öne sürer. İndirgemeci materyalizmin bir türevi olan işlevselcilik ise zihin durumlarının beynin işlevleri tarafından yaratıldığını belirtir. İşlevselci bakış açısına göre her zihinsel duruma tekabül eden belli bir nöronal aktivasyon veya birçok bağlantılı nöronal aktivasyon söz konusudur.

A tipi indirgemecilik, tasarlanabilirlik ve bilgi argümanlarını, bir cümlenin doğruluğu veya yanlışlığı gösterilebilir olmadığı sürece o cümlenin anlamlı olmadığını savunması üstünden değerlendirir. Bir önermenin düşünülebilir olması o önermenin olabilir olması anlamına gelmez.

Chalmers'ın sınıflandırmasında A tipi indirgemeci bir yaklaşım sergileyen Dennett, zor problemin bir illüzyon olduğunu ve bilincin kolay problemleri çözüldüğünde bilincin zor probleminin anlamını yitireceğini düşünen filozoflardandır. Dennett, zor problemin içeriklerinin neden zihinsel temsillerimiz olduğu ve fiziksel bir sistemin nasıl zihinsel temsiller yarattığı sorularını, DNA'nın keşfinden önceki yaşamı anlamaya çalışan vitalistlerin yaşamın özünün fiziksel bir kuramla çözülemeyeceğini iddia etmelerine benzetir. Günümüzden bir vitalistin hayali düşünce deneyiyle DNA'ya sahip ama yaşamayan bir kedi düşünmekle, sinir sistemi bilinçli bir sinir sistemiyle aynı olup, bilinçsiz bir zombi düşünmek arasında fark olmadığını iddia eder. Dennett'a göre zombilerin düşünülebilir olması, günümüzde bir vitalistin "DNA, proteinler ve benzeri

şeylerin hepsi çok iyi, ama kedi gibi görünüp, kedi gibi davranan, damarlarından kan akan, hücrelerinde DNA olan ama yaşamayan bir kedi keşfettiğimi hayal edebiliyorum" demesiyle aynıdır. (Dennett, 2011, s: 329)

Chalmers, Dennett'ın zor problem ve vitalistleri karşılaştırmasının yanıltıcı olduğunu ileri sürer. Vitalistler, hayatın fiziksel bir açıklaması olamayacağını düşünmüştürler, ancak DNA'nın keşfi, metabolik aktiviteleri, adaptasyonu, büyümeyi, kısaca canlılığın fiziksel özelliklerinin çıkış noktasını açıklayarak canlıların sergilediği özelliklere bir temel sağlamış, böylece daha önce fiziksel olarak açıklanamayan biyolojik fonksiyonların açıklaması fiziksel bir metotla yapılabilmiştir. Sonuçta canlı olmak belli fonksiyonlara sahip olmak olarak tanımlandığından, DNA, vitalistlerin açıklanamayacağını düşünmüş oldukları bu fonksiyonlara bir açıklama getirmiştir. Ancak, Chalmers, bilincin zor probleminin, bilincin belli fonksiyon ve işlevlerini tanımlamaktan öte açıklamalar gerektirdiğini savunur, dolayısıyla Dennett'ın analojisini kabul etmez. (Chalmers, 2003, s: 12)

Dennett, zihinsel temsillerin duyusal algı, iç temsiller ve davranış çıktılarını kapsadığını, tüm ilişkilerin nedensellik yasalarına uygun olduğunu, dolayısıyla bilinçli bir zihinle aynı davranışlara sahip zombilerin olanaklı olamayacağını belirtir. Dennett, bilincin beynin diğer işlevlerinden ayrılabilen tek başına bir özellik olmadığını vurgulayarak, zombilerin düşünülebilir olmadığı sonucuna varır. (Dennett, 2011, s: 328-330)

Dennett, bilgi argümanının ise kötü tasarlanmış bir düşünce deneyi olduğunu ve Mary'nin renkler hakkında bilinebilecek tüm fiziksel bilgileri bilmesinin hayal edilemeyecek bir durum olduğunu ileri sürer. Dennett, Jackson'ın düşünce deneyini değiştirerek farklı bir şekilde ifade eder; "Mary'yi esir alanlar bir gün onun renkleri görmesinin vaktinin geldiğini hükmederler. Bir tuzak olarak, onun ilk renk deneyiminde sunulmak üzere, parlak mavi bir muz hazırlarlar. Mary söz konusu muza bakar ve kandırılmakta olduğunu, muzların sarı olduğunu ifade eder. Onu esir alan adamların şaşkınlıktan dili tutulur." (Dennett, 2011, s: 466) Mary, mavi bir nesnenin ve sarı bir nesnenin sinir sistemi üstündeki tüm etkilerini bildiğinden muzun mavi olduğunu anlamıştır.

Dennett, bir şeyin tüm fiziksel bilgisinin aslında o şeyin özellikleri hakkında tüm bilgiyi içerdiğini iddia eder, ancak ona göre tüm fiziksel bilginin ne olduğunu hayal edebilmemiz mümkün değildir. Daha önemlisi ise renklerin fiziksel dünyada var olan

fiziksel özellikler olmayışıdır. Kırmızı rengi belli bir dalga boyuna sahip enerjinin zihinde tasarlanmış fikridir, beynimizde belli dalga boylarını ifade edecek temsiller yaratılmıştır ve zihin bu temsillerinden belirli bir dalga boyuna ait olan temsili kırımızı renk olarak yorumlamaktadır. Dennett, on yedinci yüzyılda Locke'un renkler, aromalar, tatlar ve sesler gibi özellikleri ikincil nitelik olarak adlandırmış olduğunu hatırlatır. İkincil nitelikler, zihnin içindeki kendinde şeyler değillerdir, daha çok normal gözlemcilerin zihinlerindeki belli şeyleri üretmek ve uyarlamak üzere dünyadaki şeylerin güçleridirler. (Dennett, 2011, s: 431) Felsefeciler, zihinde yaratılan belli fikirleri iç nitelikler, fenomenal nitelikler, *qualia* gibi kavramlarla tanımlamış ve bu kavramların zihnin temel özelliklerini oluşturduğunu iddia etmiştirler. Ancak tüm zihinsel temsillerin sadece temsil olup, bu temsiller evrim sürecinde bazı durumları ayırt etmenin sağladığı yarar nedeniyle oluşturulmuştur. Gözlemcilerin beyinlerinin ayırt edici durumları çeşitli birincil özelliklere sahiptir ve beynin birincil özellikleri ile ikincil özellikleri ayırt edebilmektedirler.

Chalmers'a göre A tipi indirgemeci yaklaşımlar, orta düzey bir X özelliğini -örneğin zihin temsillerinin nasıl oluştuğunu- belli fonksiyonlarca açıkladıklarında, bilincin açıklamasını yapmış olduklarını iddia ederler. Bu açıklama, sistemin sahip olduğu P durumunun davranışlarla ilişkisini kurmakla aynıdır. Örneğin bilinçli bir kırmızı renk algısının, beynin P durumundaki aktivasyonlarda olduğunu belirtmek, bilincin doğasını açıklamaktansa beynin hangi durumlarının hangi zihinsel durumlara denk geldiğini göstermektir. Dolayısıyla A tipi indirgemeci yaklaşımlar bilinci açıklamamaktadır.

B tipi indirgemecilik, fiziksel ve fenomenal doğrular arasında epistemik açıdan boşluk olduğunu ama ontolojik bir boşluk olmadığını, dolayısıyla zombilerin düşünülebilse de mümkün olmadığını ve Mary'nin kırmızı hakkında fenomenolojik bilgisinin eksik olduğunu kabul eder. B tipi indirgemeciliğe göre zor problem sadece epistemik bir sorundan kaynaklanmaktadır ve fenomenal durumların fiziksel durumlarla ilişkisinin ortaya konması lazımdır. (Chalmers, 2003, s: 13-14)

Chalmers'ın B tipi indirgemeci yaklaşım olarak adlandırdığı yaklaşım Levine'ın açıklayıcı boşluk makalesinde belirttiği yaklaşımıdır. Levine, Kripke'nin beyin ve zihin arasındaki ontolojik farkını kabul etmek için beynin ve zihnin farklı olduğunun düşünülebilmesinin yeterli olmadığını öne sürmüştür. Beyin ve zihin arasında ontolojik boşluk olmadığını, ancak açıklayıcı boşluk argümanında öne sürülmüş olduğu gibi epistemolojik bir boşluk olduğunu savunmuştur. Beynin fiziksel durumları fenomenal

durumların ortaya çıkması için gereklidir. Fenomenal durumlar, fiziksel durumların ya da fiziksel işlevlerin sonucu olarak ortaya çıkar, dolayısıyla fenomenal durumlar, fiziksel durumların bir alt kümesidir. Eğer fenomenal durumları fiziksel durumlar yaratıyorsa beyin, ve zihin arasında ontolojik bir boşluk bulunmaz.

Levine, yerçekimi kanunlarının yer çekimini tamamıyla açıkladığına, yine de yer çekimi sabiti için herhangi bir açıklama yapmaksızın, yer çekimi sabitini verilmiş kabul ettiğine dikkat çeker. Yer çekimi sabitinin daha fazla açıklaması yapılamamakta, yer çekimi sabiti daha basit prensiplere indirgenememektedir. Yer çekimi sabitinin doğanın açıklanamayan kaba bir gerçeği oluşu gibi *qualia*nın da açıklanamaz bir gerçek olma ihtimali vardır. Levine'e göre *qualia*yı açıklamak adına ya önemli bazı bilgilerden yoksunuzdur ya da söz konusu bir *qualia* uzayını yaratan tüm fonksiyonel korelasyonları bulmanın ötesinde *qualia* uzayının varlığını yer çekimi sabitini doğanın bir gerçeği olarak kabul ettiğimiz gibi kabul etmemiz gerekir. (Levine, 1983, s: 257-358) Bilinç sadece yüksek derecede organize olmuş sistemlerin bir özelliği gibi gözükmektedir. Levine, *qualia* uzayının üst düzeyde organize olmuş beynin fiziksel durumlarının sonucuna indirgenmesinin veya insan beyninin rutin bir hesap sistemi olduğunu kabul etmenin insanın akıl yürütme prensiplerine ters olduğunu belirtmiştir.

Chalmers, *qualia* uzayı ile beynin fiziksel durumları arasında bir ilişki olduğunu kabul ettiğimiz taktirde qualia uzayının beynin fiziksel durumlarının bir sonucu olduğunu kabul ederek belli fiziksel durumlarla *qualia* uzayı arasındaki korelasyonu veren temel bir denklem kurulmasının mümkün olduğunu kabul eder. Ancak bilimde temel yasaların her zaman farklı özellikleri birbirine bağladığını ve fiziksel durumlar ile bilinç arasındaki epistemik bağlantı temel bir yasa olarak kabul edildiğinde, bilincin fiziksel olmayan nitelikleri olduğu sonucu çıkacağını belirtir. (Chalmers, 2003, s: 17)

Chalmers'a göre bu durumda B tipi indirgemecilik, bilincin fiziksel olmadığını savunan indirgemeci olmayan modellere çökmektedir. B tipi indirgemeci modellere entegre bilgi teorisi örnek verilebilir. Entegre bilgi teorisinde beynin belli fiziksel durumlarının, beynin bütünde fark yaratan nöronal farklılaşmaların, bilgi taşıdığı, bilinçli durumların bilgi ile alakalı olduğunu savunur. Dolayısıyla entegre bilgi teorisi fiziksel bir durumu fiziksel olmayan bir kavram olan bilgi kavramı ile ilişkilendirmektedir.

C tipi indirgemecilik ise bilincin zor probleminin prensipte çözülebilir olması gerektiğini ama şu anki limitlerimizin bu problemi çözmeye yetmediğini savunur. C tipi materyalistler, nasıl ki Einstein'in kuramlarına dek enerji ve maddenin arasındaki

ilişkinin anlaşılması fiziğin limitlerinin ötesinde kalmışsa, aynı şekilde günümüzde bilincin zor probleminin bilim ve felsefenin limitlerinin ötesinde bir problem olduğunu belirtirler. Ancak bu düşünce kalıbının içinde bilinci açıklamak için yapılacak her girişim, eğer indirgemeci olacaksa, A veya B tipi indirgemeciliğe çökmek veya indirgemeci olmayan modelleri kapsamak durumunda kalacaktır. (Chalmers, 2003, s: 27)

İndirgemeci olmayan yaklaşımlar, kendilerini, materyalist açıklamalara karşı gelen argümanların doğru olması durumunda fiziksel dünyanın özelliklerini genişletmemiz gerektiğini fikrinden yola çıkarak temellendiririrler. Kütle, enerji, uzay-zaman kavramlarının fiziksel dünyanın temel kavramları oluşu ve bu kavramları açıklamak için daha derine inmeye çalışılmayışı gibi, bilincin de dünyanın fiziksel olmayan temel bir kavramı olarak ele alınabileceğini düşünürler. (Chalmers, 2003, s: 28)

Bu fikirde iki olasılık vardır; bilinç fiziksel olmayan dünyanın temel bir kavramıdır veya bilinç fiziksel olmayan dünyanın temel kavramlarından oluşmuştur. İki durumda da açıklanması gereken bilincin fiziksel dünya ile ilişkisidir. Chalmers indirgemeci olmayan modelleri bu ilişki içinde üç sınıfta inceler.

İlk indirgemeci olmayan yaklaşım, Descartes'ın ruh beden ilişkisinin de içinde bulunduğu etkileşimciliktir. Fiziksel sistemlerin neden sonuç ilişkisiyle kendi içinde kapalı bir sistem olmadığını ve fenomenal durumların fiziksel dünyayla etkileşim içinde bulunduğunu savunur. Ancak bu görüş bilimle bir uzlaşı sağlayamadığından, içinde birçok sorunu barındırır. (Chalmers, 2003, s: 29)

İkinci indirgemeci olmayan yaklaşım ise ontolojik olarak fiziksel özellikler ile fenomenal özelliklerin ayrı olduğu ve fenomenal olanın fiziksel olana etkisi olmadığını savunan epifenomenalizmdir. Fiziksel olanın fenomenal olana yol açtığını, ancak tersinin olmadığını iddia eder. Bu yaklaşım aynı zamanda fiziksel dünyanın nedensel kapalı olduğunu kabul eder. Bu yaklaşım doğru ise fenomenal durumların fiziksel durumlara bir etkisi yoktur. Ancak bu yaklaşım bilincin açıklamasına katkıda bulunabilecek bir yaklaşım gibi gözükmemektedir. Ayrıca evrimin neden daha bilinçli türlere doğru yol aldığı sorusunun cevabının daha karmaşık fiziksel sistemlerin evrilirken bilincin onlara eşlik ettiği şeklinde açıklanması tatmin edici bir açıklama olmaktan uzaktır. (Chalmers, 2003, s: 33-35)

Son indirgemeci olmayan yaklaşım, bilincin temel fiziksel birimlerin içkin özellikleri tarafından oluşturulduğunu iddia eden f-monizmdir. Bu yaklaşımda fenomenal

özellikler fiziksel gerçekliğin temelinde yer alır. Fizik temel parçacıkları temel kabul eder ve onları daha fazla irdelemez, aynı zamanda temel parçacıkların içsel özellikleri hakkında bir bilgi vermez. Bu yaklaşım kendi içinde iki metafizik soru barındırır: 1) Temel parçacıkların içkin özelikleri nedir? 2) Fenomenal özellikler temel parçacıklara nasıl entegre edilebilir? Bu sorulara rağmen, Chalmers f-monizimin hem indirgemeci, hem indirgemeci olmayan yaklaşımlar arasında en gelecek vadeden yaklaşım olduğunu savunmuştur. (Chalmers, 2003, s: 35-39) Ancak eğer bilinç temel parçacıkların sahip olduğu içkin bir özelliğin sonucu olarak ortaya çıkan bir kavramsa bilincin temelini oluşturan içkin özelliklerin temel parçacıkların davranış ve etkileşimlerinde bir etkisi olmaması f-monizm açısından ciddi bir sorun teşkil etmektedir. Bu durumda f-monizm ya epifenomenalizme çökecek ya da temel seviyede parçacıkların bilinçle alakalı içkin özelliklerinin ihmal edilebilir olduğunu ve bu içkin özelliklerin ancak belli bir karmaşıklık düzeyinde gözlemlenebilir fenomenler ortaya çıkardığını savunmak durumunda kalacaktır.

İndirgemeci ve indirgemeci olmayan yaklaşımların hepsi kendi içlerinde birçok sorun barındırmakta ve günümüzde bilincin zor problemi gizemini korumaktadır.

Günümüzde indirgemeci modeller tam anlamıyla bilincin sorunlarını açıkça çözememiş olsa dahi, zihnin çalışma prensipleri ve bilincin neden ve nasıl ortaya çıktığı konusunda oldukça kapsamlı cevaplar vermektedirler. Bilincin tüm problemleri indirgemeci modellerle çözülemeyecek dahi olsa, bu modeller bilincin sorunları hakkında ayrıntılı detaylar sunmaktadır. Crick'in belirttiği gibi fizik, kimya ve biyolojinin tüm başarılarının temelini oluşturan indirgemeci yaklaşımları terk etmek için elimizde çok güçlü nedenlerimiz olmalıdır. (Crick, 1997, s: 9)

## 3.4 ÖZNEL VE NESNEL BAKIŞ AÇILARI

Nagel üçüncü şahıs bakış açısıyla bilincin araştırılmasının imkansızlığına dikkat çekmiş ve bilincin indirgemeci yöntemlerle anlaşılamayacağını, bilinci yaratan mekanizmalardan ontolojik farklılıklarını vurgulayarak savunmuştur.

Nagel, "Yarasa Olmak Nasıl Bir Şeydir" başlıklı makalesinde birinci tekil şahsın deneyimlerinin üçüncü şahıslar tarafından erişilemez olduğunu belirtir. Nagel, zihin beden problemini zorlaştıran şeyin bilinç olduğunu söyleyerek başladığı makalesinde, zihnin indirgemeci yaklaşımlarının zihin ve beyin ilişkisine ışık tutma ihtimalinin düşük

olduğunu, zihinsel bir olgunun fiziksel doğasını açıklamak için elimizde hiçbir fikrimiz olmadığını belirmiştir. Bir organizmanın en küçük şekilde de olsa bilinçli bir deneyim yaşamasının, o organizma olmak gibi bir şeyin mevcut olduğu anlamına geldiğini söyleyen Nagel'e göre, "O organizma olmak gibi bir şey varsa o organizma bilinçli durumlara sahiptir." (Nagel, 1974, s: 147) Bir organizma olmak gibi bir durumu yaratan ise deneyimlerin öznel nitelikleridir. Nagel, zihin durumlarını açıklayan herhangi bir yaklaşımın öznel deneyimlerin niteliklerini de içinde bulundurması gerektiğini, ancak bilimin nesnel, bir başka deyişle üçüncü şahıs bakış açısıyla bu durumlara ulaşmasının imkansız olduğunu iddia eder. Öznel deneyimlerin nitelikleri işlevsel durumları açıklayan sistemlerle analiz edilemez,- çünkü Nagel öznel deneyimlere sahip olmaksızın insan gibi davranan robotların olası olduğunu düşünür.

Nagel, öznellik ve bakış açısı arasındaki ilişkiye örnek vermek, öznel niteliklerin önemini açığa kavuşturmak, öznel ve nesnel deneyimlerin arasındaki farkı net bir şekilde ortaya koymak için bir yarasa olmanın nasıl bir şey olduğunu sorgular. Öncelikle Nagel memeli bir canlı olan yarasanın öznel deneyimleri olduğunu kabul eder. Yarasalar dış dünyayı öncelikli olarak sonar seslerle algılamaktadır. Bir yarasa çevresini algılamak için yüksek frekanslı çığlıklar atar ve seslerin nesnelere çarparak nasıl yansıdığını dışarıdaki nesnelerin konumlarıyla ilişkilendirir. Yarasaların sonar sistemleri dış nesneler hakkında insanın görme sistemini işlevini yerine getirmektedir. Yarasa sonarı bir algı türü olmakla birlikte insanların sahip olduğu hiçbir algı türlerine benzemez. Sonar algısını tecrübe hatta hayal edebileceğimizi düşünmek bile imkansızdır. Nagel, bir yarasanın iç deneyimini, öznel dünyasını hangi yöntemlerle anlayabileceğimizi dikkatte almamız gerektiğini belirtir. Bir yarasa olmanın nasıl bir şey olmak olduğunu hayal etmemiz işe yaramayacaktır. Hedef, bir yarasa olmanın nasıl bir şey olduğunu anlamaktır. Nagel, bir yarasa olmanın nasıl bir şey olduğuna dair şematik bir fikirden daha fazlasını oluşturmamızın mümkün olmadığını belirtir. Aynı şekilde doğuştan kör ve sağır bir insanın olmanın nasıl bir şey olduğunu da anlamamız mümkün değildir, aynı şekilde kör bir insanın görme duyusuna sahip bir insan olmanın nasıl bir şey olduğunu anlamasın mümkün olmayışı gibi. Nagel öznel deneyimi ve bir şey gibi olmanın nasıl bir şey olduğunu düşünmemize imkan verecek kavramlar ve bir teori geliştirmemiz gerektiğini ancak doğamızın sınırlarının bizi böyle bir bilgiyi edinmekten sonsuza dek mahrum edebileceğini belirtir. (Nagel, 1974, s: 162)

Nagel fenomenolojik olguların tamamen nesnel olduğu bir boyuta dikkat çeker, bir insan bir başka insanın deneyiminin niteliğini dile getirebilir. Ancak bu nesnellik sadece atıf yapılan nesneye onun bakış açısını kullanabilecek kadar benzeyen biri için mümkündür. Eğer öznenin bakış açısına tamamen uzaksak ya da öznenin bakış açısı erişilmezse deneyimin gözlemlenebilen nesnel niteliğinin ne anlama gelebileceğini anlamak zordur. Bir yarasanın sonarının çalışma prensiplerini inceleyebiliriz ancak bu inceleme bir yarasa olmanın öznel deneyimini tamamen ortadan kaldıracaktır. Nagel, yarasa olmanın öznel deneyimi ortadan kalkacak olursa yarasa olmanın nasıl olduğuna dair geriye bir şey kalmayacağını belirtir. Nagel, aynı şekilde bir Marslının beyinlerimizi inceleyerek insan olmanın nasıl bir şey olduğunu asla anlamayacağına dikkat çeker. (Nagel, 1974, s: 156) Bilimin indirgemeci metodu nesnel olmaya ve şeylerin gerçek doğası hakkında daha doğru ve kesin bir bakış açısına ulaşmaya yönelik bir girişimdir. Nagel, indirgemeci yolu izlemeyi seçmemizin mümkün olduğunu belirtir. Bunun nedeni dış dünya hakkında düşünmek için kullandığımız kavram ve düşüncelerin başlangıçta bizim algısal donanımımızı içeren bir bakış açısını uygulamaya koymasına rağmen, bu kavram ve düşüncelerin onların ötesindeki şeylerden bahsetmek için kullanmamızdır. Eğer zihinsel süreçler aslında fiziksel süreçler ise belli fiziksel süreçlerden geçecek bir şey var demektir. (Nagel, 1974, s: 159) Ancak Nagel, şu an için indirgemeci yaklaşımın bilinci incelemek için nasıl bir metot geliştirebileceği hakkında hiçbir fikrimiz olmadığını, indirgemeci metotla öznel deneyimleri anlamamız için teorik bir çerçeveye ihtiyacımız olduğunu ancak böyle bir teoriyi nasıl yaratacağımız hakkında henüz bir bilgimiz olmadığını belirtir. (Nagel, 1974, s: 162)

Dolayısıyla Nagel, bir deneyimin, belli bir an ve belli bir durumda ancak birinci tekil şahıs tarafından yaşanabilecek bir durum olduğunu ve bilinç probleminin, içerdiği öznellik nedeniyle nesnel olarak açıklanma olasılığı olmayabileceğini belirtmiştir. Nagel, eğer indirgemeci modeller öznel yapıyı içeremiyorsa bu modellerin eksik kalacağını vurgular. Ancak Dennett bu görüşe katılmayacak ve öznel yapıyı içermeyen modellerinde epistemolojik olarak bilinci açıklayacağını savunacaktır.

# 4. BÖLÜM

# DENNETT'IN ZİHİN FELSEFESİ VE YAPAY ZEKA

Dennett güçlü yapay zekanın epistemolojik bir problem olan çerçeve probleminin çözümüyle mümkün olduğunu savunmaktadır. Fiziksel duruşu benimseyen Dennett, bilincin tamamen beynin fiziksel süreçlerinin bir ürünü olduğunu iddia eder. Zihnimizin beynimiz olduğunu düşünmek zordur ama imkansız değildir. (Dennett, 2011, s: 26) Dennett'ın yönelimsel duruşu benimseyerek bilincin özelliklerini araştırmasının nedeni ise zihnin fiziksel süreçlerin sonucu beliren bir kavram olduğu iddiasını temellendirmektir.

## 4.1 DÜALİZM ELEŞTİRİSİ

Dennett'e göre ikicilik umutsuz bir bakış açısıdır ve ikiciliği savunan filozoflar esas olarak bilincin değerini kaybetmekten korkarlar. Bazı düşünürler hala bilincin tek değerli şey olduğu dışında her şeyi reddetmektedir çünkü bilinç kavramı, ahlak, özgür irade, aşk, acı gibi insanlığın değer verdiği ve değerlerini yarattığı kavramlarla yakından alakalıdır. Dolayısıyla bilincin bilimsel bir açıklamasının yapılması bilinçle alakalı kavramlara olan bakış açımızı da etkileyecek ve değiştirecektir. Bilinç kavramı sadece bilimle açıklanırsa ahlak ve özgür irade kavramlarının nasıl bir değişime uğrayacağı, eğer bilinçli bir deneyim, bir şekilde salt bir hareket sorununa indirgenirse aşk, acı, rüya ve zevk kavramlarının nasıl açıklanacağı, ya da bilinçli insanlar sadece maddi nesneler olursa, onların haklı veya haksız olmasının bizim için ne ifade edeceğini önemli sorulardır. Bilincin ana özelliklerinin sadece tanık olunan duygular olmadığını, bilincin aynı zamanda beğenmenin, sevmenin gerçekleştiği yer olduğunu belirten Dennett, bilincin herhangi bir şeyin neden önemli olabileceğine karar veren hakem olduğunu vurgular. Bilincin ortaya çıkış nedeninin ise hayatta kalmamız üzerine düşünme ihtiyacımız olabileceğini belirtir. (Dennett, 2011, s: 35)

Bilinci Pandora'nın kutusuna benzeten Dennett, bu nedenle insanların da bilimin işgaline karşı koyma garantisi veren felsefi doktrinlere dayandıklarını, bu doktrinlerin güçlü ve zayıf yönleri olmasına karşın iyi temellendirilmiş olduklarını belirtir. Bilimin bilince verilen değeri değiştirme ihtimali vardır, Wordsworth'ün söylediği gibi, bazen

insanlar bilimin inceleme amacıyla katletmesinden korkarlar. Bilincin bilim tarafından açıklanmasının belki insanlık için kayıpları olacaktır, ancak Dennett, bilimsel bir bilinç açıklamasının kaybettireceklerinin, iyi bir bilinç kuramının hem bilimsel ve toplumsal hem de kuramsal ve ahlaki bakımdan sağlayacaklarıyla telafi edileceği görüşündedir. (Dennett, 2011, s: 33)

Dennett, ikiciliğin umutsuzluğunu Descartes'ın ikiciliğini eleştirerek vurgular. Zihin ve beden iki ayrı töz olsa da birbirlerini etkilemek zorundadırlar. Zihin, duyu organlarının elde ettiği bilgilere ulaşmak zorunda, beden zihnin uygun gördüğü davranışları yerine getirmek zorundadır. Descartes beyindeki epifiz bezi sayesinde iki tözün etkileşime girdiğini iddia etmiştir. Varsayım gereği zihinsel töz fiziksel değildir, hiçbir fiziksel ve kütlesel enerji onunla ilişkilendiremez. Dolayısıyla fiziksel olmayan bu tözün, hücrelerle nasıl etkileşime girdiği açıklanamaz. Dennett'e göre zihinsel tözün en çekici özelliği, sonsuza dek bilimi kendisinden uzak tutacak olmasıdır. (Dennett 2011, s: 49)

Bilinç, bilimin son gizemlerinden biridir. Evrende uzay, zaman, kütle çekimi gibi birçok gizem bulunduğu inkar edilemez ancak bilim, bu gizemleri çözmek adına yöntemler geliştirmiştir, bu gizemler üstüne nasıl düşüneceği hakkında fikri vardır. Ancak bilim henüz bilinci incelemek için net bir yöntem geliştirememiştir, ayrıca bilinç üstüne nasıl düşüneceğini de bulamamıştır. Bilimin genel yöntemi "böl ve keşfet"tir. Dennett bilimin böl ve keşfet yönteminin beyni alt sistemlere böldüğünü, her alt sistemin yapı ve fonksiyonlarını birbirinden ayrı incelendiğini belirtir. Alt sistemlere gösterilen dikkatinse genelde kuramsal miyopluğa neden olduğunu ekler. Dennett, ayrıca Descartes'in ikiciliğinin günümüzde materyalist bakış açısına da nüfuz etmiş olduğu görüşündedir. (Dennett, 2011, s: 55) Bilimin miyopluğu, Descartes'ın ikiciliğinin etkileriyle birleştiğinde bilim adamların bilincin beyinde bir nokta veya noktalara indirgeme çabasının ortaya çıktığını iddia eden Dennett, alt sistemlerin işlevlerinin bir araya gelerek bir noktada bilinci oluşturduğu varsayılan modellerin tümünü Kartezyen tiyatro olarak adlandırır**.** Kartezyen tiyatroda bilinç, maddi süreçler dahilinde merkezi bir konumda varlığını sürdürür, bu düşüncede bilinç beyinde olup biteni seyreden bir seyirci, bir *homunculus* gibidir. Dolayısıyla Dennett, bilimde bilinç araştırmalarının en yaygın kabul edilmiş olan yöntemin "bilincin nöronal korelasyonlar"ını bulmanın, bilinci beynin bir noktasındaki bir işlevle tanımlamaya çalıştığını düşünmektedir. "Garip bir şekilde nörobilimcilerde ikicilere benzerler, çünkü bazı şeylerin bilinçte temsilini göstererek, işin yükünü zihne yıkarlar." (Dennett, 2011, s: 298) Dolayısıyla Dennett'a göre bilincin nöronal korelasyonlarını bulma düşüncesi, Kartezyen tiyatro

düşüncesinin bir ürünüdür. Dennett'a göre bilinç beynin tek bir noktasında değildir ve zihnimizde fenomenlerin oluşturulduğu bir Kartezyen tiyatro yoktur. Bilinci herhangi açıklama girişimi ilk olarak Kartezyen tiyatro düşüncesinden, beynin aktivitelerini gözlemleyen bir *homuculunus*un varlığından kurtulmalıdır.

Kartezyen tiyatro ya da beynin içindeki bilincin içeriklerinin sunulduğu bir *homuculunus* düşüncesi, bilincin zor probleminin tanımlanmasına neden olur. Dennett, Chalmers'in zor probleminin bir illüzyon olduğu görüşündedir. Dennett'a göre zor problem yoktur, zor problem bilinci açıklamaya çalışan yaklaşımların yanlışlığından dolayı varmış gibi gözüken bir problemdir. Bu yaklaşımların en temel hatası *qualia*yı zihnin temel bir özelliği olarak kabul ediyor olmalarıdır. (Dennett, 2017, s: 3) Aynen Kartezyen tiyatro düşüncesi gibi *qualia*nın zihnin temel özelliği olduğu düşüncesi de Descartes'ın özellik ikiciliğine dayanmaktadır. Descartes, her şeyden şüphe etmiş ama kendi bilinçli deneyimlerinin içeriklerinden, *qualia*sının varlığından şüphe etmemiştir.

Dennett, *qualia*nın birinci şahıs açısından varlığını sorgulamaz. *Qualia*, dünyadaki şeylerin özelliklerinin bizim zihnimizdeki temsilleridir. Zihnimizde dünyaya ait temsiller vardır. Dennett, dünyadaki cisimlerin algısal özelliklerinin, renklerin, seslerin, tatların, dokuların zihnindeki varlığını, ancak paranın varlığını inkar ettiği kadar inkar etmektedir. *Qualia*da aynen para gibi bir temsildir, zihnimizde temsiller vardır ama bu temsiller bizim düşündüğümüz şeyler değildirler. (Dennett, 2011, s: 35) Bilinci oluşturan şeylerin tümünü -*qualia*nın içeriklerini- fenomenoloji bahçesi olarak adlandıran Dennett, fenomenoloji bahçesini üç ana bölüme ayırır: (1) sesler, görüntüler, kokular, kayganlık veya kaşıntı gibi dış dünya deneyimleri; (2) fantezi imgeleri, iç görüler, gündüz düşleri, hatırlamalar ve ön seziler gibi iç dünyanın saf deneyimleri; (3) bedensel hislerden, açlık, susuzluk hissine, öfke, neşe, utanç, şehvet, şaşkınlık gibi duygulanımlar. (Dennett, 2011, s: 55)

Dennett'a göre bilinci oluşturan bu şeyleri ve bilinç akışını inceleyecek uzman fenomenologlar yoktur. Elemeci materyalist bir yaklaşımla Dennett, bilinci oluşturan şeylerin insanların düşündüklerinden çok farklı olduğunu ve eski terimlerin kullanılmaması gerektiğini savunur. "Fenomenolojinin Yokluğu" başlıklı makalesinde bunu bir analojiyle açıklar. Hayvan diye bir şey olmadığını iddia eden bir adamı yanılgısıyla yüzleşmek için hayvanat bahçesine götürmenin mümkün olduğunu, ama bu adamın hayvanların kürklerle kaplı birer robot olduğunu iddia etmeye devam edebileceğini belirtir. Dennett'a göre bu deliliktir ama bu adam hayvanların doğasına

ait devrimci bir fikre de sahip olabilir. Zoologlar hayvanlar konusunda uzmandır ve hayvanat bahçelerinde kendi uzmanlık alanlarını insanlara tanıtırlar. Eğer zoologlar bu adamın fikirlerini haklı bulurlarsa, kendi keşiflerinde de adamın fikrini kullanmanın bir yolunu bulacaktırlar. Ancak fenomenoloji söz konusu olduğunda bilincin aslında farklı bir şey olduğu iddia edilecek olursa, bu fikri değerlendirecek uzman bir fenomenolog yoktur. Bilinç ve bilinç akışı ile ilgili düşünürken hayvanat bahçesini gezen bir ziyaretçi gibi düşünürüz. Dennett, bilinç akışı içindeki acıların, korkuların, düşlerin ve zihinsel imgelerin, kısacası fenomenolojik deneyimlerimiz anlamına gelen *quali*anın bizim düşündüğümüz şey olmadığını, bu şeylerin doğasını bilmediğimizi ve bu şeyler için yeni sözcükler bulunması gerektiğini savunur. (Dennett, 1979, s: 94-95)

Dennett, *quali*anın gerçekte ne olup, ne olmadığını “Quining Qualia” makalesinde inceler. *Quining* kelimesi gerçek bir şeyin varlığını veya önemini kararlılıkla reddetmek anlamına gelmektedir. (Dennett, 1988, s: 1) Dennett, gerçek olan her şeyin belli özelliklere sahip olduğunu belirtir. Kendisi bilinçli deneyimin içeriklerinin, dolayısıyla *quali*anın varlığını reddetmediğinden, makalesinde *quali*anın gerçek özellikleri olup, olmadığını çeşitli düşünce deneyleriyle araştırmıştır.

Dennett, ilk olarak *qualia* için yapılmış genel tanımlardan *quali*anın dört özelliğini belirler. Deneyimi oluşturan her elementin kendine has değişmez, parçalanamaz özellikleri *quali*aya içkin özellikleridir. Dolayısıyla *quali*ayı oluşturan temsillerin içkin özellikleri vardır. Örneğin görsel bir deneyimdeki kırmızı başka bir zihin içeriğindeki renk ya da temsil ile veya değişik temsillerin birleşimi ile açıklanamaz. Her fenomenolojik deneyim, deneyimleyen birinci sahsın deneyimidir, dolayısıyla sahsa özeldir. Bir deneyimin temsilleri, *quali*anın içerikleri ne kadar üst düzey bir dil ile açıklanırsa açıklansın, o deneyimin içeriğini tam olarak ifade etmek imkansızdır, dolayısıyla *quali*anın içeriği dil ile aktarılamaz, ifade edilemez. Ayrıca bir deneyimin nitelikleri bilinçte dolaysız olarak anlaşılırdır. Dennett, *quali*ya affedilen bu dört özelliğin sezgisel ve daha önemlisi yanlış özellikler olduğunu düşünmektedir. (Dennett, 1988, s: 1-2)

Dennett, zevkle buharda pişmiş karnabahar yiyen birini tahayyül eder. Karnabahar yiyen kişinin karnabahardan aldığı tat kesinlikle onun karnabahardan aldığı tatla farklı olmalıdır. Dahası Dennett, her yiyeceği her zaman aynı tada sahipmiş gibi hissetmediğini, değişik zamanlarda kahvenin tadının ona farklı geldiğini belirtir. Dolayısıyla *quali*anın tat uzayında belli bir yiyeceğin tadı zamana göre değişmektedir. Dolayısıyla *quali*ayı oluşturan *quale*ler birbirilerinden, zamandan ve değişik fiziksel

etkilerden bağımsız değildirler. Dennett, bir şarap tatma makinesinin tasarlandığında ve bu makine şarap aromalarını incelikle ayırt ettiğinde makinenin deneyimine *qualia*nın eşlik etmeyeceğini belirtir. Ancak bu makine her zaman şarapları doğru şekilde ayırt edip, sınıflandıracaktır. Dolayısıyla bu makine, bir şarap tadıcısının tat alma, deneyimlediği tattan keyif alma gibi birinci şahıs deneyimlerinin sonuçlarıyla eşdeğer sonuçlar elde edebilmektedir.

Dennett'ın bir başka düşünce deneyinde bir kişi bir beyin cerrahının şakasına maruz kalır. Beyin cerrahı, kurban uyurken kurbanın beyninde birtakım değişiklikler yapar ve kurban uyandığında tüm objelerin renklerinin değişmiş olduğunu fark eder; çimleri mavi, gökyüzünü ise yeşil görmektedir. İlk olarak bu kişi *qualia* uzayında renklerin yerlerinin değiştiğini fark edecektir ancak iç gözlemle *qualia* uzayındaki renklerin nasıl değişmiş olduğunu anlayamayacaktır. Cerrah kurbanın optik sinirlerinin yerini değiştirmiş olabileceği gibi, kurbanın hafızasında tanımlanmış renkleri de değiştirmiş olabilir. Bu noktada Dennett, *qualia* uzayında renk spektrumun değişmesinin iki fiziksel nedeni olabileceğini ancak kurbanın kendi *qualia* uzayına doğrudan ve özel erişimi olmadığından dolayı kendi *qualia*sındaki değişimin nedeni hakkında bir fikri olmadığını belirtir. Kurban *qualia*sında bir değişim olduğunun farkındadır ve bunun nedenini çözemediği için üçüncü bir şahsa danışması, doktora gitmeye karar vermesi muhtemeldir. Dolayısıyla *qualia* iç gözlemler ile anlaşılabilir ve birinci şahıs tarafından direkt erişilebilir değildir.

Dennett'ın sunduğu bir başka düşünce deneyi ise "kahve tadıcıları"dır. Altı yıl boyunca her sabah aynı kahveyi tatmış olan Chase, bir sabah kahvenin tadının aynı olduğunu ancak artık ağız tadının değiştiğini ve aynı kahveyi sevmediğini belirtir. Diğer kahve tadıcısı Sanborn'da altı yıl boyunca her sabah aynı kahveyi tatmıştır ve bir sabah kahvenin tadının ve kendi aldığı tadın aynı olduğunu ancak tat reseptörlerindeki bir problem nedeniyle kahvenin tadından keyif almadığını belirtir. Dennett, Sanborn ve Chase'in kendi *qualia* durumlarında yanılma olasılığı olduğunu ve ikisinde kendileri hakkında yaptıkları açıklamalara güvenemeyeceğimizi iddia eder. Halbuki Chase ve Sanburn'un kendi *qualia*larına dolaysız ve anlık erişimi olmuş olsa onların yaptığı açıklamaları kabul etmemiz gerekir.

Dennett, yukarıdaki düşünce deneylerinin *qualia*nın kişiye özel ve kişinin direkt erişimi olan bir yapısı olmadığını gösterdiğini iddia eder. *Qualia*nın içerikleri fiziksel etkilerle değiştirilebildiği gibi, deneyimin sahibi birinci şahıslar, *qualia* uzayındaki bir değişikliğin nedenini iç gözlemle açıklayamamaktadır. Ayrıca eğer *qualia*yı zihnin

ontolojisine dahil edeceksek, qualia'nın kararlı ve tanımlanabilen özellikleri olması gerektiğini savunur. Düşünce deneyleri *qualia*nın kararlı tanımlarının olmadığını, aynı birinci şahsın dahi *qualia* uzayındaki bir özelliği değişik zamanlarda farklı tanımladığını da göstermiştir. Dennett'a göre zihinsel temsillerimizin varlığı inkar edilemez. Ama temsillerdeki epistemik ilişkiler dışsal epistemik olanlarla tam anlamıyla aynı olmadığı için, *qualia* zihnin doğasına ait temel bir özellik değildir. (Dennett, 1998, s: 12)

Dennett'ın bir başka örneği ise "deneyimli bira içicisi"dir. İlk kez biranın tadına bakan birisinin biranın tadını sevmesi pek muhtemel değildir ancak bira içtikçe bira içicileri birayı sevmeye başlarlar. Biranın tadı değişmediğine göre biranın tadının kötü iken zamanla güzel olması ilişkisel bir durumdur ve bira içicisinin zihinsel durumu, bira içerken yaşadıkları, hafızası biranın tadı hakkındaki görüşlerini değiştirmiştir. Halbuki içkin bir özelliğin ilişkisel olmaması, dolayısıyla gözlemciye göre değişmemesi gerekir. (Dennett, 1998, s: 13)

Dennett, sonuç olarak *qualia* kavramının evrendeki diğer temel kavramlar gibi tanımlanmış, kendine has değişmeyen özellikleri olan, başka kavramlarla açıklanamayan, temel bir kavram olmadığı belirtir. *Qualia* kavramı aynen DNA'nın keşfinden önce yaşamı tanımlama çabası içindeki biyologların kullandığı ve yaşamın özü olduğu düşünülen "*elen vital*" kavramı gibidir. Ancak filozof ve nörobilimciler *qualia*dan bahsederken neden bahsettiklerini biliyor gibi konuştukları ve *qualia*nın inkar edilemez temel bir zihinsel özellik olduğunu düşündükleri için *qualia*nın varlığı zor problemi yaratmaktadır. (Dennett, 2017, s: 3) Bilinci açıklayan bir model Kartezyen tiyatro düşüncesinden kurtulmalı, dolayısıyla beynin içinde bilincin içeriklerinin sunulduğu bir öznenin olduğu düşüncesinden vazgeçmelidir. Bilinç beynin bir parçasına sunulan bir kavram değildir, beynin tüm aktivitesinin sonucudur. Dennett, beynin içindeki öznenin saf dışı bırakılması gerektiğinde ısrarlıdır; beyinde kimse yoktur ancak beynin içindeki özneyi saf dışı bırakmak, zor problemi görmezden gelmek olarak yorumlanmaktadır. Dennett, zor problemin zaman içinde kolay problemlerin bir parçasına dönüşeceğine inanır.

Dennett, zor problemin ve bilincin içeriklerini temsil eden *qualia*nın bir illüzyon olduğunu savunur. Sonuçta zihnimiz dünyayı kendisi oluşturmaktadır. Gördüğümüz dünya gerçek dünya değil, zihnimizde oluşturduğumuz bir dünyadır. Beyin dış dünyadan aldığı sınırlı bilgiyi, önceden oluşturduğu temsilleri kullandığı yukarıdan aşağı işlemeli süreçlerle birleştirerek zihinsel temsillerini oluşturur. *Qualia*nın içerikleri

bir insanın kendi zihinsel durumları hakkında düşünceye ya da iç gözleme dayanan inançların tasarlanmış nesneleridirler. Dış dünyanın temsillerinin sunulduğu bir Kartezyen tiyatro düşüncesi illaki sahneyi seyreden bir özne gerektirir, beynin içinde bir öznenin varlığı direkt olarak zor problemi doğurmaktadır. Dennett, fenomenolojinin bilincin doğasını araştırmak için doğru bir yöntem olmadığını belirtir ve bilinci incelerken kullanılacak bilimsel bir yöntem geliştirmeye çalışır. Bilinç birinci tekil şahsa ait bir kavram olduğundan, bilincin bilimin bir konusu olabilmesi için nötr bir yönteme ihtiyaç vardır. Fenomenologlar tarafından benimsenen bakış açısı birinci tekil şahıs bakış açısıdır. Dennett, birinci tekil şahıs bakış açısını bilinci açıklamak için yetersiz bulur, bilinç içeriklerini açıklamak için fenomenolojinin ötesine geçmek zorunda olduğumuzu belirtir. (Dennett, 2011, s: 81) Dennett bilimsel bir duruş için üçüncü tekil şahıs bakış açısına sahip olmamız gerektiğini vurgular. Başlangıçta fenomenolojiyi verili kabul eden Dennett, birinci tekil şahıs deneyimlerine ulaşmak için hetero-fenomenoloji yöntemini benimser.

Dennett, bilinci açıklayan en güzel metaforik çizimlerden birinin Newyorker dergisinin kapak fotoğrafı olduğunu belirtir. (Dennett, 2017, s: 1) Bir sergide tabloya bakan adamın bilinç akışını canlandıran bu çizimde, ressamın adını hatırlamaya çalışan adamın bilinç akışı karikatürize edilmiştir. Ressamın ismi Barok kelimesini çağrıştırır; Barok kelimesi başka kavramları çağrıştırır, bilinç akışında çağrışımlarla oluşan kavramların yanı sıra renkler, biçimler, muhtemelen çizimde ifade edilemeyen kokular, sesler, hisler bulunmaktadır. Dennett, bir bilinç modelinin çizimdeki adamın o anki zihinsel süreçlerini açıklaması gerektiğini düşünür, ancak adamın bilinç akışında bulunan her nesnenin, adamın bir inancının ürünü olduğunu ekler. Bilinç akışındaki nesnelerin hepsi adamın inançları doğrultusunda adamın zihninde yaratılmış temsillerdir. Bu temsilleri yaratan mekanizma beyindir, beynin aktivitelerinden bir kısmına adamın erişimi vardır ve onun rapor etmesiyle zihinsel temsilere üçüncü şahıslarında erişimi olacaktır. Dennett, bu noktada adamın zihin akışındaki temsilerin sadece beynin bir aktivitesi olduğunu ve temsillerin gerçeklikle direkt ilişkisi olmayışının önemli olduğunu belirtir. Metaforlar her zaman gerçekle paralel bir durumu aktarmazlar. Hetero-fenomenolojik yöntem rasyonel taşıyıcının betimlediği temsilleri incelemelidir ancak bu temsillerin gerçekliği yansıttığını ya da doğru olduğunu kabul etmemelidir, adamın zihindeki temsiller Noel baba kadar gerçek dışı olabilir. Kırmızı bir elma ile kırmızı bir elmanın zihindeki temsilini bağdaştırmakla, kırmızı elmanın beyindeki temsillerini yaratan, kırmızı ve yuvarlak *qualia*yı yaratan nöronal aktiviteyi bağdaştırmak farklıdır. Çünkü zihindeki kırmızı ve yuvarlak *qualia*yı

yaratan nesne ne yuvarlak, ne de kırmızıdır. (Dennett, 2017, s: 2) Zihnimizdeki yuvarlak şekil görme korteksinde retinadan topografik bağlantılarla gelen sinyallerle oluşturulmuş bir aktivasyon örüntüsü, kırmızı ise V4 alanındaki belli nöronların aktivasyonudur. *Qualia*daki kırmızı elma ise bu aktivasyonun sonucu yaratılan zihinsel bir temsil, bir illüzyondur.

## 4.2 ÇOKLU TASLAKLAR MODELİ

Dennett, “çoklu taslaklar modeli” olarak bilinen bilinç modelini oluştururken fenomenolojiyi verili kabul eder. Çoklu taslaklar modelinin sağlaması gereken en kritik koşul ise Kartezyen tiyatro yanılgısına düşmemektir; dolayısıyla modelde bir sahne, sahnede öznenin dış dünyayı ve kendi iç dünyasını anlaması için sergilenen bir oyun olmamalıdır. Dennett bilincin bir sahneden daha çok beyindeki nöronların ani ve dinamik ateşlenme örüntüleriyle açıklanması gerektiğini düşünmektedir. Bu örüntülerde zihnin temsillerini yaratan asıl süreçler duyusal girdileri ipuçları olarak alarak yukarıdan aşağı akışlarla tüm olası temsilleri oluşturan süreçlerdir. Dolayısıyla Dennett’ın çoklu taslaklar modelinde bilinç akışından daha çok zamanla rastgele değişen ama sürekliliği olan nöronal aktivasyonlarla oluşturulan bir bilinç içeriği söz konusudur.

Çoklu taslaklar modelinde Dennett, ilk olarak nörobilimin bulgularından beynin yüksek derecede paralel bir işlemci olduğunu, beynin modüler bilgi işleme yapısını ve beyindeki çoğu sürecin bilinç dışı gerçekleştiğini, bu süreçlerden ufak bir kısmının bilinçli içeriği oluşturduğunu verili kabul eder. Duyusal uyaranların işlenmesi sırasında beynin hiyerarşik bilgi işleme yapılarının işlevlerini yerine getirirken, beynin bir tahmin etme makinesi gibi çalışarak aldığı duyusal uyaranın sunduğu ipuçlarından yukarıdan aşağı bilgi akışıyla uyaranın temsilinin yaratıldığı da Dennett’ın modelinde kabul edilmiştir. Dennett, özellikle beynin yukarıdan aşağı mekanizmalar ile kendi temsillerini nasıl oluşturduğuna, duyusal algıdaki boşlukları nasıl doldurduğuna dikkat çeker. Kendinden bir örnek veren Dennett, Belotto’nun Dresden manzarası olarak bilinen geniş bir perspektifle resmedilmiş tablosuna gerekli uzaklıktan baktığında uzak bir köprüde resmedilmiş insanları ayrıntılarıyla gördüğünü düşündüğünü, ancak tabloya yaklaştığında uzaktan ayrıntılı insan figürleri olarak görünen şekillerin ustaca kullanılmış ancak yakından anlam ifade etmeyen fırça boyaları olduğunu fark ettiğini belirtir. (Dennett, 2003, s: 12) Görsel sistem fırça darbelerini analiz etmemiş,

resmedilenlerin insan olduğunu belirlediği anda yukarıdan aşağı bilgi akışıyla fırça darbelerini ayrıntılı insan figürleri çizimleri olarak doldurmuştur. Dennett, yukarıdan aşağı süreçlerin duyusal temsilin yeniden temsili olduğunu belirtir. Yeniden temsil, Dennett'a göre ikinci dereceden bir temsil olan öznenin bilinçli deneyimini oluşturur.

Görsel bir uyarandan yansıyan fotonlar retinaya düştüğü anda bilgi retinada işlenmeye başlar. Dennett, işlenmekte olan bilginin bilinçli bir içerik oluşturup oluşturmayacağının ilk anda bilinmesinin imkansız olduğunu belirterek, belli bir süreci tamamlayan bilginin bilinçli içerik oluşturması gerekmediğini iddia eder ve bir içeriğin bilinçli içeriğe dönüşmesinin koşullarının ne olduğunu sorgular. İlk olarak lokal işlenen bir bilgi parçası, çevresinde işlenmekte olan bilgilere göre daha fazla fark yaratmalı ve belli bir süre kararlı olarak sesini duyurmayı başarmalıdır. Eğer söz konusu bilgi yeterince kararlı durumda sesini duyurmazsa bu bilgi bilinç dışında işlenmeye devam edecek ama yine de değişik modüllere ulaşacaktır. Dennett'a göre, kartezyen tiyatro, bir sürecin ilk olarak tamamlandığını ve sonucunda bilinçli bir içeriğin ortaya çıktığını iddia etmektedir. (Dennett, 2011, s: 131-132) Dennett'ın çoklu taslaklar modelinde ise işlenmekte olan her bir bilgi bir taslaktır ve birçok taslakla rekabet halindedir. Dennett'ın çoklu taslaklar modelinde, nektar küpünün iki farklı açıdan görünüşü istatistiki olarak aynı şansa sahip iki taslağın rekabeti olarak yorumlanır.

Bir taslağın bilinçli içerik yaratması bir anda gerçekleşmez; taslak yaratılmaya başlandığı an itibariyle, taslak çevresindeki nöronal devrelerden ve yukarı aşağı süreçlerden gelen bilgiler ile revize edilmeye de başlamıştır. Bu revizyonlar Orwellvari veya Stalinvari revizyonlardır. “Orwell revizyon”u yaşanmış bir olayın sonradan değiştirilerek güncellenmesi, “Stalin revizyon”u ise sahte bir temel üstüne, yanlış bir duyusal algı üstüne, bilginin işlenmesi anlamına gelir. Örneğin bir köşede dururken uzun saçlı bir kadının koşarak geçtiğini gördünüz. Bir saniye sonra gözlüklü bir kadın geçsin ve hafızanızı kirletsin. Uzun saçlı kadınla ilgili detaylar sorulduğunda, Orwellvari bir revizyonla, onun gözlüklü olduğunu söylemeniz muhtemeldir. Belleğin Orwell revizyonu ya da algılama esnasında oluşan Stalinvari revizyonlardan Dennett, bilinçli bir deneyimin son bulduğu bir zamansal nokta olmadığını, bilinçli deneyimlerin belli bir anda oluşturulmuş birçok taslağı içerebileceğini, taslakların başka taslaklardan etkilenerek değişebileceğini ve beynin boşlukları kendisinin doldurmaya eğilimli olduğunu öngörür. (Dennett, 2011, s: 140-146)

Çoklu taslaklar modelinde bilinçli deneyimin içeriği sınırları çizilmiş tek bir anlatıdan oluşmaz. Paralel işlenmekte olan birçok, birbiriyle çakışan, revize edilmiş içerikler

birbirini etkiler. Bilinçli bir deneyim esnasında beyindeki paralel bilgi işleme sürekli devam eder; denekten, bilinçli deneyimi farklı zamanlarda sorgulandığında, farklı cevaplar alınması muhtemeldir. Bilinçli deneyimin içeriği, içeriğin ne zaman ve nasıl sorgulandığı ile alakalıdır; çünkü bilinçli içerik devam eden dinamik bir sürecin ürünüdür. Dolayısıyla Dennett, bilinci akış olarak algıladığımızı, ancak aslında bilinç içeriklerinin devamlı değişen, bir anda diğerlerini bastırarak öne geçen taslakların bir bütünü olduğunu savunur. Beyinde, alt modüllerin hesaplamalarını tamamlayıp, tamamlanmış alt düzey hesaplar birleştirerek bilinçli bir sahne kurulan zamansal bir bitiş noktası bulunmaz. Nesnel olarak beyinde fark yaratan süreçler anı tayin edilebilse bile, bu sıra öznel deneyimin sırasını belirlememektedir. (Dennett, 2011, s: 161-164) Beynin gerçek zamanlı davranışları kontrol etmesi gerekir, beyindeki temsillerin işlevi gerçek zamanda davranış kontrolüdür. Zamanlama algısal işlemlerin başında içeriği belirleyebilir, ancak temsiller uygun davranışın kontrollüne katkı sağladığı sürece, hangi sıralama içerisinde meydana geldiğinin bir önemi yoktur. Temsil etmenin zamanlaması içeriği belirler. Zamanlamanın temsili ise bir başka durumdur. (Dennett, 2011, s: 171-182)

Dennett'ın ayrıca bilinç içeriğinin tüm aktivasyonlarının geçmişe yönelik bir şekilde iz bırakmasıyla ortaya çıktığını iddia eder. Bu iddiasının arkasında ise Libet deneylerinin çıkarımları vardır. Libet, 1980'li yıllarda yaptığı bir takım deneylerde hareketin, hareketin bilinçli kararından önce geldiğini gösteren bir dizi deney yapmıştır. Deneylerin bir versiyonunda deneğin beyin aktivitesi kaydedilirken, denekten istediği bir anda bileğini bükmesi, bileğini bükmeye karar verdiği anı da ayrıca rapor etmesi istenmiştir. Beklenen sonuç, beyin aktivitesinin karar verme ve hareket sırasıyla gerçekleşmesiyken, deneylerde ilk olarak hareketi başlatan aktivasyon, daha sonra ise bilinçli kararın aktivasyonu gözlemlenmiştir. Libet, bilinçli istem ve eylemleri gerçekten başlatan beyin olayları arasında 300-500 ms aralığında bir fark olduğunu belirtmiştir. (Libet, 1999, s: 50-51) Beyinde her olayın belli bir uzamsal mekânsal yeri vardır. Dennett'a göre bileğimizi ne zaman bükmeye karar verdiğimizi sormak başlangıç anından bitiş anına dek ne yapacağımızın veya uyaranın miktarının bilincinde olduğumuz anlamına gelir. Halbuki bilinç tüm beynin aktivitelerinin bir sonucu olarak ortaya çıkar. Uyaranın beyinde işlenmesi, uyaranın taslağının, o an beynin tasarladığı diğer taslakların önüne geçmesi noktasal bir anda ve beynin bir noktasında gerçekleşmez. (Dennett, 2011, s: 197-198)

Dennett, “Hala Bilinci mi Açıklıyoruz” başlıklı makalesinde kendi çoklu taslaklar modeli ile temelleri Baars tarafından atılmış, Deheane ve Naccache tarafından geliştirilen kapsamlı çalışma alanı modelinin birçok noktada uyumlu olduğunu belirmiştir. (Dennett, 2001, s: 222) Dennett, verili bir zamanda bir modülde kararlı ve fark yaratan süreçlerin, beyinde daha geniş bir ölçekte yankılanarak diğer modüllerde fark yarattığını, global olarak söz konusu bilginin tüm sistemi etkilediğini kabul eder. Nöronal çalışma alanı modelinde basit modüllerde fark yaratan süreçlerin, kapsamlı çalışma alanı nöronlarıyla beynin diğer modülerinde yankılanması ve amplifiye edilmesi söz konusudur. Dennett’ın çoklu taslaklar modeli ile kapsamlı nöronal çalışma modeli arasındaki fark, Dennett’ın bu noktada değişik modüllerin tüm fark yaratan süreçlerinin bilinçli deneyimin bir parçasını oluşturduğunu, her modülde fark yaratan süreçlerin ileri bir işleme tabi tutulmadığını iddia etmesidir. Fark yaratan bir modül, bilinçli deneyimin öncülü olarak başka modüller tarafından değerlendirilmez. Fark yaratan süreçler bir noktada, bir sahnede, ya da bir üst temsilde birleştirilerek bilinçli deneyimi yaratmaz. (Dennett, 2001, s: 235-236) Dennett’ın bilinci akıştan daha çok anlık nöronal aktivasyonlarla tanımladığı göz önünde tutulduğunda fark yaratan modüllerdeki bilginin diğer modüllerde yankılanması ve belli modüllerdeki yukarıdan aşağı revizyonlarla bir bilinç sahnesi oluşturur. Aktivasyonların birbirini etkileyerek revize etmesi, geçmiş ve geleceğe yönelik revizyonları da kapsar, bilinç sahnelerinin art arda gelmesi ise bilinç akışını yaratır.

Dennett, her modülde fark yaratan ve kararlı duruma geçen aktivasyonların kendilerinin bilincin içeriğini oluşturduğunu ve üst düzey bilişsel sistemlerin bilince sunulan bilgiyi kullanmaktan daha çok aynı şekilde paralel şekilde fark yaratan, kararlı durumdaki bu aktivasyonları ara bir adım olmaksızın kullandığını düşünmektedir. Dolayısıyla Dennett’a göre bir modülde diğer aktivasyonlarla yarışarak fark yaratan aktivasyonlar, bilinçli deneyim oluşturmak için diğer fark yaratan aktivasyonlara ihtiyaç duymazlar; bir an için fark yaratırlar ve bir sonraki fark yaratan aktivasyon tarafından bastırılana dek bilincin içeriğini oluşturma şansını elde ederler. Eğer söz konusu aktivasyon dil ile ifade edilecekse süreç içinde korteksin değişik bölgelerine yayılmış olduğu düşünülen dil ile ilgili nöronal devreler bu bilgiyi anlık kullanır, önce belli bir fark yaratan sürecin tamamlanmasını beklemezler.

Özet olarak Dennett, bilinci açıklayacak çoklu taslaklar modelini oluştururken nörobilim araştırmalarının bulgularını verili kabul eder. Beyin ağırlıklı olarak paralel işleyen modüler bir yapıya sahiptir ve alt modülerin iyi tanımlanmış görevleri bulunur.

Zihin dünyayı, evrimsel olarak geliştirmiş olduğu bir biçimde, dış dünyadan aldığı ipuçlarını yorumlayarak kendi içinde yeniden yaratır. Zihnin temsilleri ile dış dünyanın gerçekliği arasında doğrudan bir bağlantı kurulamaz. Genel olarak beyinde oluşturulan temsillerde yukarıdan aşağı süreçler baskındır ve her yeni temsil bir önceki temsili değiştirir. Dennett, beyindeki süreçlerin ancak küçük bir kısmının bilinçli içeriklere dönüştüğünü de kabul eder. Dennett'ın çoklu taslaklar modelinde bilincin içeriği farklı modüllerde tanımlanmıştır. Değişik modüller bilincin içeriğini oluştururken yukarıdan aşağı süreçlerle duyusal algıdan gelen ipuçlarını kullanarak birçok taslak oluştururlar. Kapsamlı çalışma alanı modelinde olduğu gibi her modülde oluşturulan çoklu taslaklar diğer modüllerin taslaklarıyla iletişim halindedir. Oluşturulan birçok taslaktan hangisinin bilincin içeriğini oluşturacağı ise hangi taslağın belli bir süre içinde kararlı kalarak diğer taslaklardan daha fazla fark yaratacağı ile belirlenir. Dennett'ın çoklu taslaklar modeli Kartezyen tiyatro düşüncesini barındırmaz. Dennett, söz konusu bir bilinç içeriğinin tamamlanıp, sunulduğu bir sahne fikrine şiddetle karşı çıkmıştır. Aynı şekilde Dennett'ın modelinde bilinç akışından daha çok devamlı değişen ve revize edilen bilinç içerikleri bulunur. Dennett'ın modelinde beyinde paralel bir şekilde aynı anda gerçekleşen bir aktivite söz konusudur. Paralel aktivite sırasında her modül kendi görevini yerine getirirken birçok mikro yargı oluşturur. Bu yargılar hem beynin yapısında zaman içinde oluşmuş temsillerle, hem de söz konusu aktivitenin hemen öncesindeki aktivitelerle bağlantılı şekilde oluşurlar. Modüller içinde beliren içerikler veya yargılar modüller arasındaki iletişimle dinamik zihin taslakları oluştururlar. Oluşan her taslak dinamik olarak eklemelere, düzeltmelere, üstüne yazmalara maruz kalır. Daha sonra bu taslaklar birbirleriyle rekabet ederek tüm beyinde kalıcı bir hakimiyet kurmaya çalışırlar. Bu hakimiyet bir sonraki adımda ne düşünüleceğini, neye dikkat edileceğini veya neye odaklanılacağını da belirler. Ancak bir taslağın ne zaman bilinçli olacağı bir muammadır. Taslaklardan hangisinin bilinçli deneyimi oluşturacağı baştan bilinemediği için Dennett, bir temsilin nesnel ve öznel zamanlamasını farklı değerlendirir. Belirli herhangi bir zaman sürecinde bilincinde olduğumuz taslak, sürece dair başlattığımız araştırmadan bağımsız olarak tamamlanmaz, devamlı bir revizyon süreci devam eder, dolayısıyla kurallara uygun kabul edilecek tek bir versiyon da yoktur. Sonuçta bilinç akışı olarak gözlemlenen bilinçli deneyim süre gelen tüm bu aktivitelerin geçmişe yönelik olarak beyinde iz bırakmasından ortaya çıkar.

## 4.3 KÜLTÜRLE TAMAMLANAN BİLİNÇ MODELİ

Çoklu taslaklar modeli, beynin yapısı ve çalışma prensiplerinden yola çıkararak bilinçli deneyimin bir kartezyen tiyatronun varlığını kabul etmeksizin nasıl oluşturulacağını açıklar. Dennett'a göre bilinç modelinden kartezyen tiyatro elendiği taktirde, sahnenin sunulacağı bir özne de olmayacağından, zor problem bir illüzyona dönüşür. Ancak Dennett, çoklu taslaklar modelinin bilinci açıklamakta yetersiz olduğunu belirtir. Dennett'a göre bilincin iki birleşeni vardır. İlk birleşen bilincin biyolojik temeli olan beyindir. Dennett bilincin bir diğer birleşeninin ise kültür olduğunu belirtir. Çoklu taslaklar modeli, bilincin beyinle olan ilişkisini açıklamaktadır ancak Dennett bilincin birçok özelliğinin kültürle doğmuş olduğunu ve bilincin biyolojik temeli beyin olsa da kültürün bilinci oluşturmuş olduğunu iddia eder. (Dennett, 1991, s: 35)

Dennett, bilinç kavramını nörobilim araştırmacılarının birincil bilinç olarak adlandırdığı bilinç kavramı gibi tanımlamamaktadır. Dennett'ta bilinç, fiziksel bir temelde vuku bulan ancak kültürle tetiklenen ve kültürü de yapısına katan bir tanıma sahiptir. Dennett'ın iki birleşeni olan bilinç tanımıyla hayvanların veya bebeklerin bilincinden bahsetmek mümkün değildir. Dennett, insan bilincinin hem evrim sürecini değiştirdiğini, hem de insan bilincinin evrim sürecinden bağımsız anlaşılmasının mümkün olmadığını düşünerek, bilincin evrimde hangi koşullarda ortaya çıkmış olduğunu araştırır.

Yaşam ağacına baktığımızda ilk önce tek hücrelileri görürüz. Daha sonra bu tek hücreli canlılar ökaryotik devrimle tek hücreli güçlerini birleştirerek çok hücreli ilk canlıları yaratmışlardır. Diğer süreçlerle kıyaslandığında çok daha kısa bir zaman süresinde ise kambriyum patlaması yaşanmış, bu dönemin ardından çeşitlilik hiç olmadığı kadar hızlı artarak memeliler ve *homo sapiens* ortaya çıkmıştır. *Homo sapiens* binlerce yıl önce ilk ortaya çıktığında dünyanın biokütlesinin sadece binde birini oluştururken, bugün yüzde doksan sekizini oluşturur. Bu büyüme dünyanın tanık olduğu en hızlı gelişimdir. Bu gelişim buzul çağında olmuş ve insanlar, buzul çağının başından sonuna diğer yakın akrabalarına fark atarak beyinlerini dört kat büyütmeyi başarmışlardır. Dennett bu büyük değişimin sadece genlerdeki evrimle açıklanamayacağını ve bu gelişimin nedeninin farklı olduğunu söyleyerek değişimin temelinde insan bilincinin ortaya çıkışını yerleştirir. (Dennett, 2011, s: 217)

*Bakteriden Bach'a* başlıklı kitabında Dennett, Gaudi'nin ünlü *La Sagrada Familia* kilisesi ile termit yuvalarını kıyaslar. Termit yuvaları şekil ve tasarım olarak Gaudi'nin

kilisesiyle neredeyse aynıdır. Termit yuvası, ne yaptığından bihaber olan milyonlarca termit tarafından inşa edilirken kilise dahiyane, planlı bir tasarımdır. Dennett, Gaudi'nin beyninin de yarı özerk ve ne yaptıkları hakkında tek başlarına fikirleri olmayan milyonlarca nöronlardan oluştuğunu belirtir. Dennett, termit yuvası ve insan beyni arasındaki tartışmasız benzerliğin ötesinde, termit yuvası ve insan beyninin arasındaki farkın insan bilincini açıklayacak olan anahtar olduğunu ileri sürer. Termit yuvası ve insan beyni arasındaki ilk fark termit kolonisindeki termitler belirlenmiş katı kurallar ile hareket etmesi, insan beyninin ise yukarıdan aşağı işleme mekanizmalarına sahip olmasıdır. (Dennett, 2009, s: 2) Termit yuvası ve insan beyni arasındaki ikinci fark ise insan beyni düşünce aletlerine sahiptir. Dennett, evrim sürecinin bir aşamasında düşünme aletlerinin insan beynine yerleştirilmiş olduğunu belirtir. Düşünme aletleri zekanın idrak edilmesini sağlayan aletlerdir. (Dennett, 2017, s: 16)

Dennett, başlangıçta aklın olmadığını, sadece etkenlerin var olduğunu belirtir. Derken ilk yaşam formlarının temeli olacak kendini kopyalayan kromozonlar ortaya çıkmıştır. Bu kromozonların kendilerini kopyalarken bir amacı olmamıştır ama insan bakış açısıyla yaşamın kökeninde kendilerini kopyalamayı amaç edinmiş protein zincirleri ortaya çıkmıştır. Eğer bu eşleyiciler dünyadaki entropiye karşın var olmaya devam edeceklerse belli şartları sağlamak zorundaydılar ama diğer taraftan evrimin milyonlarca mutasyonundan entropiye karşı var olabilen eşleyiciler var olmaya devam etmişlerdir. Varlığını koruyan eşleyiciler, kötü şeylerden kaçınıp, iyi şeyleri isteyerek dengeleşimlerini devam ettiren eşleyiciler olmuşlardır. Bu eşleyiciler dünyadaki olayları kabaca faydalı, sakıncalı ve nötr olarak bölümlendirecek bir bakış açıları varmış gibi görmüş ve ilk anlarından itibaren iyi olana yönelme ve kötü olandan sakınma davranışları oluşturacak mekanizmalar inşa etmeye başlamışlardır. Eşleyicilerin karşılaştıkları ilk problem, amaçlarını nasıl anlayacaklarını öğrenmek ve bizatihi kendi varoluşlarının ortaya çıkması için amaçlarına göre eylemekti. (Dennett, 2011, s: 205) Basit organizmaların bulduğu çözüm ise çevrelerine dokunmaktı; basit fiziksel temaslarla çevrelerindeki iyi ve kötü faktörleri algılıyor ve çekilme veya yakınlaşma tepkileri veriyorlardı. Damasio, Humphrey gibi birçok bilim adamı ilk tek hücreli organizmaların çevreleri tarafından yaratılan yaklaş veya geri çekil mekanizmalarının bilincin en önemli özelliklerinin temelini oluşturduğunu düşünürler.

Tek hücreli prokaryotlar dünyada uzun bir süre egemenliklerini ilan etmiş ve çeşitlenmişlerdir. Ökoryatik devrim, bu tek hücrelerin birbirleriyle birlik oluşturarak

içsel sistemleri olan hücreler meydana getirmesidir. Ökoryatik devrimde bir noktada, bir prokaryot bir başka prokaryotun bedenine girerek ilk ökoryatı oluşturmuştur. Bu dönemde, günümüzde çekirdekli hücrelerin hepsinde bulunan ve hücreye enerji sağlayan mitokondriler, içine girdikleri prokaryotlarla iş birliği yapmışlardır. Ökaryotlar hücreler, bizim hücrelerimizde de olduğu gibi, yapılarında iki DNA barındırırlar; biri mitokondri DNA'sı, diğeri ise ev sahibi olan prokaryot DNA'sıdır. Çekirdekli hücreler yaşam ağacında ilk bağımsızlığını ilan eden canlılar olmuşturlar. Çekirdekli hücrelerin oluşturdukları içsel sistemler, gelişmiş organizmaların vücutlarındaki ilk sistemlerin temelini atmıştır. Yaşamı yönetmek ve yönlendirmek için oluşan bu sistemlerde ödül ve ceza ile ilgili moleküller, ileri türlerin duygularını yönlendiren moleküllerin öncülleri olmuşturlar.

Yaşam ağacında bir sonraki büyük değişim ise kambriyen patlaması olarak adlandırılır. Kambriyen patlamasında dünya üstündeki çeşitlilik daha önce hiç olmadığı kadar artmış, çok hücreli canlılar tüm dünyaya yayılmıştırlar. Çok hücreli canlıların oluşumunu ise içsel sistemleri olan ökaryotik hücrelerin çeşitliliğinin ve çok yönlülüğünün ulaştığı muazzam çeşitlilik sağlanmıştır. Ökaryotik hücrelerin çeşitliliği, çok hücreli canlılarda organ hücrelerinin özelleşmesini mümkün kılmıştır.

Çok hücreli organizmalarda bulunan hücreler ile tek hücreli organizmaların hücreleri arasındaki en önemli fark tek hücrelilerin kendi başlarının çaresine bakması gerekirken çok hücreli organizmaları oluşturan hücrelerin karmaşık ve çeşitli bireylerden oluşan topluluklarda yaşamalarıdır. Çok hücreli organizmalar karmaşıklaştıkça bu organizmalarını yaşamlarını idame ettirmelerini sağlayacak basit sinir sistemleri sahneye çıkmıştır. İlk basit sinir sistemlerinin zihnin temellerini atacak şekilde gelişmesi ise, Maclver'ın iddiasına göre, omurgalıların denizlerden karaya çıkmasıyla gerçekleşmiştir. Maclver, karada av ve avcı ilişkisinin geleceği ön görme yetisinin önemini arttırdığını, geleceği öngörme yetisinin ise beyinlerin zihinleri oluşturmasında kilit bir yeri olduğunu öne sürmüştür. (Maclver, 2017, s: 2375) Geleceği öngörme yetisi hayal gücünün ilk adımıdır. Sinir sistemleri geleceği ön görme yeteneği, gelecekte ne olacağına uygun davranma yeteneği geliştirerek kendilerine büyük bir avantaj sağlamıştırlar.

Doğal seçilim kat ettiği yolu nasıl kat ettiğini anlatamaz ama doğal seçilim ile tasarlanan sistemler ile zeki mühendisler tarafından tasarlanan sistemler arasında derin farklar vardır. (Dennett 2011, s: 206) Mühendisler tasarımlarını belli bir amaca yönelik ve aynı zamanda beklenmedik yan etki ve etkileşimleri bertaraf edecek şekilde

tasarlar. Evrim ise tasarımlarını tamamen amaçsız tesadüfler ve doğal seçilimin eleme mekanizmalarıyla oluşturur. Evrim masraflı bir mühendisliktir, dünyada bulunan türlerin kat ve katı çocuk sahibi olamadan yok olmaya mahkum bırakılmıştırlar. Evrimin tasarımlarında asla çalışan eski bir tasarım kullanımdan kalkmaz, her yenilik eski tasarımların üstüne monte edilir. Evrim ileriyi hiç görmediği için öngörülemez yan etkileri düşünmez, onları önlemeye çalışmayarak tasarımlarını dener. İki veya daha fazla ilişkisiz işlevsel sistem bir araya gelip, sürprizler üretebilirler, bu sürprizlerden zeki tasarımcıların tasarımlarında hiç bulunmayan tek ögelerin çoklu işlevleri ortaya çıkar. Çok işlevli tasarımlar aynı zamandan üst üste eklenen tasarımların da bir yan ürünüdür. (Dennett, 2011, s: 207)

Crick, evrimin evrendeki en zeki tasarımcı olduğunu belirtmiştir. Doğada evrimin tasarımlarındaki zekayı sadece canlıların vücutlarında değil, misal olarak termit kolonisinin inşa etmiş olduğu Gaudi tarzındaki termit yuvasında da gözlemleriz. Evrimin dahice tasarımları zeki bir tasarımcının varlığına ihtiyaç olmaksızın, belli kurallarla, olasılıklarla ve acımasız elemelerle ortaya çıkmıştır.

Darwin'in en temel fikri, evrimin amaçsız ve tamamen kör güçler tarafından yöneltilen bir süreç oluşudur. Dennett, Robert Mckenzie Beveley'in Darwin'in Türlerin Kökeni üstüne yazmış olduğu sert eleştirinin, Darwin'in fikirlerini en iyi şekilde yakalayan yorum olduğunu belirtir: "Başa çıkmamız gereken teori açısından zanaatkar mutlak cehalettir; ki tüm sistemin temel ilkesinin MÜKEMMEL VE GÜZEL BİR MAKİNE YARATMAK İÇİN MAKİNENİN NASIL YAPILDIĞINI ÖĞRENMENİN GEREKLİ OLMAMASI olduğunu ileri sürmektedir. Bu iddia dikkatli bir incelemeyle ve sıkıştırılmış bir yapıyla teorinin temel manasını ve Mr. Darwin'in kastettiği şeyi birkaç kelimeyle ifade eder; Darwin akıl yürütmeyi tuhaf bir şekilde tersine çevirerek, MUTKAL CEHALETİN, mutlak bilgeliğin ve yaratıcı becerinin tüm niteliklerine sahip olduğunu düşünüyor gözükmektedir." (Dennett 2017, s: 64) Darwin kesinlikle akıl yürütmeyi tersine çevirmiş ve zeki tasarımların zeki bir tasarımcıya ihtiyaç duymaksızın nasıl ortaya çıktığını göstermiştir. Darwin'in tersine akıl yürütmesiyle aynı şekilde tersine akıl yürütmüş olan bir başka bilim insanı ise Alan Turing'tir; Turing, mükemmel bir hesaplama makinesi olmak için, aritmetiğin ne olduğunu bilmenin bir ön koşul olmadığını Turing makinesiyle göstermiştir. (Dennett, 2009, s: 570)

Darwin ve Turing'in ortak olarak vurguladığı fikir ise idraktan yoksun bir zekanın varlığının mümkün olduğudur. Dennett, zihnin özelliklerinin neden değil, sonuç olarak görülmesi gerektiğini vurgular. İnsanların zihnine kazınmış geleneksel fikirlerin

temelinde zeki bir tasarımcının erek ve anlamla zihni yarattığı düşüncesi yanlış zihin modelleri ile kendini göstermektedir. Dennett, evrenin maddesel bir evren olduğunu ve zihnin çok yeni bir buluş olduğunu düşünmektedir. Termitler zeki tasarımcılar değildir ama insanlar zeki tasarımcılardır ve insan zihninin ortaya çıkışı, dünyanın yaşına göre çok yakın gelecekte gerçekleşmiş bir olaydır.

Dennett, yaşam ağacındaki en belirgin dönüşümün dünyadaki ilk zeki tasarımcı olan *homo sapiens*in ortaya çıkışı olduğunu belirtir. MacCready, *homo sapiens*in tarım toplulukları oluşturmaya başladığı zaman, yaklaşık on bin yıl önce *homo sapiens* ve *homosapiens*in sahip olduğu omurgalı hayvanların dünyadaki omurgalı biokütlenin sadece binde birini oluştururken, günümüzde bu oranın yüzde doksan sekiz olduğunu hesaplamıştır. “MacCready patlaması” olarak bilinen bu durum, tek hücrelerden günümüze dünya üzerindeki en büyük değişimdir. MacCready patlamasını sadece genlerle açıklamak mümkün değildir. Buzul çağında *homo sapiens* tüm akrabalarına fark atarak dünyaya egemen olmuştur; *homo sapiens*in egemenliği birçok türün tükenmesini de beraberinde getirmiştir. *Homo sapiens*in gelişmiş sinir sistemi MacCready patlamasını açıklamak için gereklidir ancak yeterli değildir. Bir canlının geleceği ön görme yeteneği ne kadar esnek olursa hayatta kalma şansının o derece artacağı açıktır. Doğadaki kaotik bir değişimde basma kalıp sinir sistemlerinin şansının olmayacağı gibi ani bir değişim karşısında genetik evrimle yavaşça değişen sinir sistemlerinin de başarılı olma şansı azdır. Doğadaki kaotik değişimlerde sadece kendini yeniden tasarlayabilen esnek sinir sistemleri hayatta kalır. Bu yeniden tasarlamaya bazen öğrenme, bazen gelişme denmektedir ancak ikisi arasındaki ayrım tartışmalıdır. Genel olarak bebeklerin konuşmayı öğrenmesi gelişme, bir çocuğun iki sayıyı toplayacak yetenekler geliştirmesi ise öğrenme olarak adlandırılabilir. Esas olarak bir sinir sisteminin esnekliği, organizmanın ortamında karşılaşılan yeni koşullara, organizmanın kendini yeniden organize ederek yanıt verme yeteneğidir. (Dennett, 2011, s: 217) Doğal seçilim ve öğrenme yeteneği arasında bir ilişki kurmak mümkündür ve bu ilişki “Badwin etkisi” olarak bilinmektedir. Öğrenilen yetenekler becerilerin gelişmesini sağlayarak türün devamlılığına katkıda bulunurlar.

*Homo sapiens* beynini, akrabası olduğu primatların beyinlerinden ayırt eden en belirgin özellik, hacminin akrabalarına kıyasen dört kat daha fazla olmasıdır. Hacimdeki bu artış birdenbire olmamıştır. İki ayaklı *homo sapiensin* beyin büyüklüğü yaklaşık üç buçuk milyon yıl boyunca diğer primat beyinleriyle aynı ebada sahip

olmuş, ancak buzul çağının başlangıcından, yaklaşık iki milyon yıl önce beynin muazzam büyümesi başlamış ve bu gelişim yüz elli bin yıl önce tamamlanmıştır. Dennett, beynin gelişiminin dilin gelişiminden önce tamamlanmasının önemli olduğunu vurgular. (Dennett, 2011, s: 224) İnsan beyninin dil için gerekli koşulları doğuştan yapısında taşıdığı fikri, nörobilim tarafından da desteklenen bir fikirdir. Öğrenme işini gerçekleştiren ve doğuştan gelen bir donanım olmaksızın öğrenmek mümkün değildir. Dennett, düşünce aletlerinden yoksun saf bir beynin düşünme yetisine sahip olamayacağını vurgular; ancak söz konusu beynin de düşünce aletlerini kullanabilecek fiziksel donanıma sahip olması gerekmektedir.

Dennett, ilk zamanlar *homo sapiens*in dil kullanımının diğer primatlarda olduğu gibi değişik seslerle iletişim kurarak başladığını ancak bu dil kullanımının idraktan yoksun bir yetenek olduğunu iddia eder. Bazı kelimeler ilk olarak kavramları etiketlemek için kullanılmaya başlanmıştır. Zamanla da bazı kelimeler, evcil olmadığı halde insanlarla beraber yaşayan hayvanlar gibi insanların yaşamına girmiş ve insanlar, yabani köpekleri evcilleştirdiği gibi bu kelimeleri de evcilleştirmiştir. (Dennett, 2017, s: 197) Zamanla idrak edilen kelimeler, dil gelişimin belli bir aşamasında, bilgi aktarımı için kullanmaya başlamıştır. Dennett, seslendirmenin faydalı bilgiyi ortaya çıkarma ve paylaşma işlevi gördüğü zaman, bunun dilin evriminde bir döneme tekabül ettiğini ve dili bilgi aktarımı olarak kullanıldığı bir etkinliğe katılmanın *homo sapiens* topluluğunun bireylerine avantaj sağladığından dolayı toplulukta iletişimsel alışkanlıklar yer etmiş olmanın olası olduğunu düşünür. (Dennett, 2011, s: 229) Dennett, bu noktada öz uyarımın önemini vurgular, öz uyarımın en güzel örneklerinden biri kendi kendine konuşma pratiğidir.

Dennett, öz uyarımın ilk olarak cevapsız bir yardım çağırıcısının kendine cevap vermesiyle başlamış olabileceğini, *homo sapiens*in kültürün beraberinde kendini idrak etmesinin öz uyarımla başarıldığını belirtir. Dennett'ın bir başka iddiası ise bilgi paylaşma işlevi gören dil kullanımı idrak edilmeye başlandıkça kendine soru sorma pratiğinin, başkalarının sorularını cevaplamanın bir yan etkisi olarak ortaya çıktığıdır. (Dennett, 2011, s: 230) MacCready patlamasının açıklanmasında dil kullanımı yetisi koşuldur, ancak MacCready patlamasını meydana getiren, dil yetisinin ötesinde gerçekleşen kültürel evrimidir. Bu noktada Dennett, Dawkins'in mem kavramına başvurur. Mem, bir kültür aktarım birimi veya bir taklit birimidir. (Dennett, 2011, s: 237) Önce konuşma, sonra yazı olarak kendini gösteren insan dili, kültürel evrimin gerçekleştiği bilgi atmosferini yayarak kültürel aktarımın aracı olmuştur. (Dennett

2011, s: 415) Dawkins, memlerin genlerin analoğu olduğunu ve memlerin genler gibi bilgi taşıdığını belirtmiştir. Memler bilgi yapılarıdır ve virüslere benzetilebilirler. Virüsler canlı yapılar değillerdir, RNA zincirinden oluşurlar, yaşayan bir hücrenin içine girmeyi başardıklarında, hücrenin kendi DNA'sı yerine virüs RNA'sının kopyalanmasını sağlayarak çoğalırlar. Dennett, aynı şekilde memlerinde beyinlere yerleştiğini, çünkü onları taşıyacak fiziksel bir araca ihtiyaç duyduklarını ve beyinleri kullanarak, beyinlerden başka beyinlere aktarılarak kendilerini yaydıklarını belirtir. Dennett, memlerin var olduğunun en kesin ispatının ise kelimelerin varoluşu olduğunu belirtir. Her kelime telaffuz edilebilir bir memdir ve kelimeler insanın en önemli düşünme aletlerinin temel ögeleridir. Ancak kelimeler memlere sadece birer örnektir. Ezgiler, düşünceler, çanak çömlek yapım yolları, kısacası insandan insana aktarılan faydalı veya zararlı her türlü bilgi ve alışkanlık mem örnekleridirler. (Dennett, 1995, s: 412)

*Homo sapiens*in beyinlerinin memlerle, bilgi taşıyıcı virüslerle enfekte olmuştur. Adım adım gerçekleşen bu süreçte beyinleri memlerle enfekte olan ilk insanlar, bu yetiye aynen dil yetisine de ilk başta anlamadan sahip oldukları gibi anlamadan sahip olmuş, henüz kültürden ve ne yaptıklarından habersiz bazı alışkanlıklar geliştirmiş ve kısa zamanda bu alışkanlıklar nesilden nesle aktarılırken dilinde eşlik ettiği bir kültüre evirilmiştir. Dennett, esas olarak bilgiden oluşturulmuş memlerin insan beyinlerine kurulan sanal makineler tarafından işleniyor olmasının akla yatkın olduğunu düşünür. Bilgi soyut bir kavramdır, *bit*lerle ifade edilir, ancak bilgi somut bir yapı aracılığıyla var olmak durumundadır. Dolayısıyla memlerin insan beyninde yer bulması, beynin üstüne kurulan sanal bir makine ile mümkündür. Doğuştan sahip olduğumuz beyinlerde bulunan ve evrim baskısıyla evrimleşen özelliklerin diğer türlerde bulunmadığını belirten Dennett, bu özelliklerin kültürel aktarımla beslendiğini, dolayısıyla dilin insan beyninin ayırt edici fiziksel bir özelliğini yarattığını iddia eder. (Dennett, 2011, s: 444)

İlk zamanlar basit kelimeler, alışkanlıklar, iletişim şekilleri olarak beyinlere kurulan bu sanal makinelere her jenerasyonda yeni sanal makineler eklenmiştir. Sanal makineler gitgide karmaşıklaşmış ve insanlığın kültürlerinin oluşmasını sağlamıştır. Günümüzde hiçbir insan ne tekerliği yeniden icat etmek, ne kalkülüsü baştan yaratmak, ne de müzik aletlerini yeniden tasarlamak zorundadır. Doğduğumuz andan itibaren zihnimiz binlerce fikirle bombardıman edilir, bu fikirlerin bir kısmı beyinlerimizde yer bulur. Memler her doğan insana, doğduğu andan itibaren kültürle aktarılmaktadır. İnsanlara doğdukları anda birçok düşünme aleti hazır olarak sunulmakta ve öğretilmektedir.

İnsanlık tarihinde memler yaratan ve memleri yayan büyük bilim adamları, sanatçılar ve filozoflar vardır, ancak memler sadece büyük düşünürlerin icatı değildir, memlerin çoğunun bir yaratıcısı yoktur, çoğu düşünme aleti, örneğin dil evrimsel bir süreçle gelişmiştir. Aynı şekilde viral internet memleri de günümüzün kültürünün bir parçasıdır. Kültürün ve bilginin nesilden nesle aktarımı dünyada sadece insanların sahip olduğu bir yetenektir. Bu yetenek sayesinde insanlık birçok düşünce aleti geliştirmiş, en önemlisi kendisini içinde bulunduğu kültür ve söz konusu kültürün düşünme aletleri aracılığı ile tanımlayarak kendini idrak etme yeteneği kazanmıştır. "Biz neyiz sorusunun yanıtı, kültürün bizi ne şekilde etkilediğinde yatar." (Dennett, 2011, s: 406)

Ancak kültürlerin zamanla değişmesi bağlamında bir kültürel evrim olduğuna şüphe yoktur, eskiden kalma özelliklerin bazıları korunup biriktirilirken, bazıları kaybedilir. (Dennett, 2011, s: 412) Dennett, kültürlerin de Darwin'in doğal seçilim modeline uygun bir şekilde geliştiğini belirtir. Memler, kültürlerin genleri gibi düşünülebilir. Evrim olgusunun gerçekleşebilmesi için gerekli üç koşul, çeşitlilik (muhtelif öğelerin ortamda sürekli olarak ve bolca bulunması), kalıtım (öğeler kendi kopyalarını ya da eşlerini oluşturma yeteneğine sahiptir) ve farklılaşmış uyumluluk (bir öğenin belli bir süre zarfında üretilen kopya sayısı değişkendir ve öğenin özellikleriyle içinde bulunduğu çevrenin özelliklerine bağlıdır) olarak belirlenmiştir. (Dennett, 1995, s: 410)

Günümüzde memler hiç olmadığı kadar fazla bir çeşitliliğe ulaşmıştır; konuşma, dinleme, yazma, okuma ve bilgi aktarım teknolojileri memleri yaymış ve yayılan memler değişik zihinlerde çeşitlenmiştir. Ancak bu çok çeşitli memlerin ve kopyalarının yayılması büyük ölçüde onları taşıyan fiziksel araçlara, -bu araçlar ister beyinler, ister iletişim araçları olsun- doğrudan etki edecek seçilim kuvvetlerine bağlıdır. Zihinlerin kapasitesi sınırlı olduğundan mem seçilimi kuvvetinin bir başka yönü ise zihinlere girme ve zihinleri meşgul edebilme yetenekleridir. Ancak en sonunda ancak taşıyıcılarının gerçek anlamda yararına olan memlerin, başka bir ifadeyle toplumsal dengeleşimi sağlayacak olan memlerin hayatta kalacağı ön görülebilir. Ancak dinamik bir süreç içinde bakıldığında, genlerde olduğu gibi memlerinde illaki bir şeyin yararına olması gerekmez; eşleme konusunda başarılı olan her mem yayılır, ancak kendisine sahip olan bedenlerin akıbetleri memlerin geleceğini tayin eder.

Termit yuvası idraktan yoksun bir yeteneğin örneğidir, doğanın zeki bir tasarımcısı olmasa da biyolojik dünya evrimsel yollarla zekaya sahip olmuştur. Dennett, *homo*

*sapiens*in dünyanın hakimi olmasını sağlayanın bu yeteneğin, idrak edilmesiyle açıklanabileceğini söyler. Dennett'a göre insan bilincinin kendisi beyindeki mem etkilerinden oluşan devasa bir komplekstir. (Dennett, 2011, s: 246)

Dennett, bilinç modelini tamamlamak için zihnin çoklu taslaklar modelinin üstüne mem komplekslerini nasıl eklenmiş olabileceğini araştırır. Dennett'ın ön varsayımları şunlardır: İnsan bilinci, doğuştan gelen makine içine gömülü olmak için fazla yeni bir buluştur; insan bilincinin, büyük oranda önceki alıştırmalarla beyne aktarılan kültürel bir evrimin bir ürünüdür; insan bilincinin başarılı kurulumu, beynin esnekliğindeki sayısız mikro ayarlama ile belirlenir. (Dennett, 2011, s: 257)

Her ne kadar çoklu taslaklar modeli bilinç akışından daha çok birbiriyle bağlantılı, geçmiş ve geleceğe göre revize edilen bilinç sahnelerinden oluşsa da, bilinç akışının seri yapısından dolayı memlere ev sahipliği yapacak olan sanal makinenin Von Neuman tarzı bir seri yapılı bir makine olması gerekmektedir. Beynin paralel yapısında ne hesaplanırsa hesaplansın bilinç beynin hesaplarını aşağı yukarı sıralı bir şekilde takip eder. Eski yapay zeka mimarileri olan Von Neuman mimarileri sembolik mimarilerdir. İnsanın düşünme süreçlerini kural dizileri olarak taklit eden bu mimariler bilinç akışından yola çıkararak yukarıdan aşağı doğru dizayn edilmiştirler.

Eğer zihninize konuşma yetisi sanal makinesi kuruluysa bu makine sayesinde ilk olarak kelimeleri, sonra dilbilgisi paketlerini indirebilirsiniz. Daha sonrasında, indirip, kurmuş olduğunuz dil paketiyle erişebileceğiniz diğer düşünme aletlerinin kodlarını beyninize indirebilirsiniz. Kelimeler basit memlerdir dolayısıyla bilgi taşıyan ama aynı zamanda kendi içlerinde evrim geçiren yapılardır. Aynı zamanda kelimeler en temel düşünme aletleridir, diğer düşünme aletleri ise rakamlar, haritalar, sesler ve şekillerdir. Ancak esas üst düzey düşünme aletleri, bir konu hakkında nasıl düşünüleceğinin kurallarını içeren sezgi pompalarıdır. Dennett, erken gelişim esnasında beyine, zihnin organize edilmiş, önceden deneyimlenmiş değişik alışkanlıklarının da yüklendiğini belirtir. (Dennett, 2011, s: 258) Zihin için öğrenme aslında bilgi edinme stratejilerini öğrenmektir ki zihin alışkanlıkları stratejilerin oluşturulduğu süreçlerdir. (Dennett, 2011, s: 261) İlk baştan evrimle şekillenmiş beynin üstüne kurulu bir sanal makine sayesinde zihne indirilen her düşünme aleti, idrak kapasitesini arttıracaktır. Sanal makine, beynin fiziksel yapısına bir kez kurulduktan sonra düşünme aletlerini ve sezgi pompalarını içeren yazılımlar taklit, pekiştirme, ödül ceza gibi yöntemlerle doğumdan başlayarak insan zihinlerine yerleştirilebilirler. Dennett, memlerin insan zihnine yüklenmesinin çoğunlukla dil

yoluyla ancak aynı zamanda sözcüksüz imgelerle ve diğer veri yapıları da aracılığıyla gerçekleştirdiğini belirtir. (Dennett, 2017, s: 201-202)

Dennett, çoklu taslaklar modelinde olduğu gibi, insan beynine seri yapıdaki zihni oluşturacak olan mem yığınlarını yerleştirirken de, kartezyen tiyatro yanılgısına düşmemeye özen gösterir. Zihni yaratacak olan sanal makinenin seri bir mimarisi olmalı ancak beynin içindeki bir *homunculus* olduğu düşüncesinden uzak kalınmalıdır. Neuman mimarileri kural öbekleri ile iş yapan seri makinelerdir, beyin ise paralel yapılı bir işlemcidir. Dennett, Turing'in hesaplanabilirlik tezinden yola çıkarak, teoride paralel mimarili bir makinenin yaptığı hesapların tamamını, her paralel kanalın bir dizi de temsil edileceği seri bir makine tarafından taklit edebileceğini, aynı şekilde de eğer paralel mimari seri mimariye çevrilebiliyorsa seri bir mimarinin de paralel bir mimari tarafından taklit edilebileceğini varsayar. (Dennett, 2011, s : 252) Dolayısıyla beynin paralel işlemcisi üstüne paralel mimari de bir Neuman makinesi kurmak mümkündür. İnsanın dil yetisini oluşturanda beynin paralel yapısı üstüne kurulmuş bir sanal makinedir. Bu sanal makine ilk olarak kelimeleri, dil bilgisi kural paketlerinin yüklenmesine izin verir, dil paketlerinin yüklenmesi ise diğer birçok sanal makinenin kurulumunu mümkün kılar. Dennett, bilinçli insan beyninin, evrimin bize sağlamış olduğu aşağıdan yukarı dizayn edilmiş paralel bir makine üzerinde, verimsiz bir şekilde uygulanmış az ya da çok sıralı gerçek makineler olduğunu iddia etmektedir. (Dennett, 2011, s: 255)

Bilinç, beynin paralel yapısı üstüne kurulmuş birçok sanal makinenin çalışması ve beynin paralel yapısını şekillendirmesiyle ortaya çıkar. Sanal makinelerin beynin paralel yapısındaki modülleriyle ve geri bildirimli bağlantılarıyla kural öbeklerinden oluşan sanal makineler bu modüllerdeki bilgiyi manipüle eder, birleştirir ve kullanırlar. Dennett, bu noktada, dil içeriğinin kortekste bulunan modüllere gömülü olduğunu ve semantiği oluşturan mekanizmanın çift yönlü geri bildirim mekanizmaları olduğunu iddia etmektedir. Sanal makinenin kuralları veya komut döngüleri korteksin çağrışım ilişkileri içinde temsil edilirler. Bu şekilde Dennett, sanal makinenin yapısını kortekse dağıtarak Kartezyen tiyatro mitini modelinden uzaklaştırdığı gibi gözlemlenen bilinç akışını, çoklu taslaklar modelinde tanımlanan özelliklerini korumayı başarır. Sanal makinenin ardışıklığı onun gömülü tasarım özelliğinden daha çok uzman koalisyonların ardışıklığının sonucudur. (Dennett, 2011, s: 301) İnsan beyni evrimsel gelişimle şekillenmiştir. Beynimiz yırtıcı hayvanlardan korunmak, meyve toplamak, simaları tanımak, kavramak, bir şeyler atmak, düşen bir şeyden başını eğerek kaçmak

gibi elzem görevleri yerine getirmek için evirilmiştir. Bu yapının okuma yazma veya dil gibi yeni rollere adapte olması, dayatılan eğilimlerin bizzat tasarımın ürünü olmasındandır. Memlerden oluşan sanal makineler beynimize yüklendikçe beynin işlevselliğini adım adım değiştirerek modifiye ederler. Dennett, dil gibi seri mimaride bir bilgi paylaşım sanal makinesi ve bu sanal makine aracılığıyla yüklenmiş düşünme aletlerinden yoksun bir beynin, bilinçsiz bir beyin olduğunu düşünmektedir. Dennett, kurulan sanal makineyi Joyceçu* makine olarak adlandırılır. Joyceçu makine evrimin aşağıdan yukarı, idrak yetisinden yoksun olarak gelişmiş paralel yapısı üstüne kurulmuş bir sanal makinedir. Bu makine paralel olarak çoklu taslaklar modelindeki paralel yapılar üstüne dağıtılmıştır ve hafıza küresinin de içindedir. (Dennett, 2011, s: 257) Bilinci yaratan sanal makinelerin insan zihnine yerleştirilmesi ya da yüklenmesi için, sanal makine yüklemeye yatkın bir beyin yapısı şarttır. Muhtemelen *homo sapiens*in yakın akrabalarına oranla dört kat büyümüş beyin ebatları, sanal makinelerin yüklenebilmesi açısından yeterli işlem kapasitesi ve belleği sağlamaktadır. Söz konusu paralel işlemci beynin işletim sistemi de, Dennett'a göre korteks yapısının içine gömülmüş seri mimarili bir sanal makinedir ve bu öncül sanal makine başka sanal makine ve paketlerin yüklenmesini olanaklı kılar. Dennett, yükleme işlemlerinin hem evrimsel, hem de insan hayatı içinde kademeli olarak gerçekleştiğini belirtir. Doğumdan itibaren ilk olarak dil için gerekli yazılımlar zihne aktarılır, dilin yanı sıra kültürün içerdiği birçok mikro alışkanlık paketi değer sistemlerinin ödül ve ceza süreçleri sonucu beyne yerleştirilir. (Dennett, 2011, s: 257-259) Bu süreç içinde edinilen en önemli mikro alışkanlıklardan biri de kendi kendine konuşma ve düşünme pratikleri, bir başka ifadeyle öz uyarım alışkanlıklarıdır. Bu öz uyarım alışkanlıkları kendini izleme, kendine hatırlatma, tekrar etme gibi faaliyetlerle uzun vadeli gelecek ön görüsü oluşturulmasında da hayati öneme sahiptir. Alışkanlıklar, alışkanlıkları yaratan memler, hayat boyunca, çoğunlukla sesli iletişimle, ancak aynı zamanda yazı ile yazılımlar beyne yüklenmeye devam ederler. Eğitim kendi başına devasa boyutta mem komplekslerinin ve ileri seviyede düşünme aletlerinin beyne yüklenmesidir. Bu yüklemeleri olanaklı kılan öz uyarımlar ise aynı zamanda benliğin oluşmasını sağlar. Nörobilim alanındaki araştırmacılar kendilik bilincini otobiyografik benlik olarak adlandırarak epizodik bellekle sıkı bir ilişki içinde oluşturulduğunu öne sürerken, Dennett'a göre kendilik bilinci mem komplekslerinden yaratılmış bir illüzyondur. Benlik, dili anlamlandıran anlamdırıcıdır, anlamsa dil ile

---

* James Joyce romanlarında bilinç akışının zihinsel içeriğin kıvrımlı dizisi olarak tarif ettiğinden Dennett, zihni oluşturmak üzere kurulan sanal makineyi Joyceçu makine olarak adlandırmıştır.

vuku bulmaktadır. Dennett, Fransızca konuşan bir zihnin, Japonca konuşan bir zihinden, okur yazar olan bir zihin ise cahil bir zihinden farklı olduğunu, bu değişik zihinlerin kendi benlik bilinçlerini de farklı anlamlandıracağını belirtir. (Dennett, 2011, s: 243)

Dennett'ın insan beyninin yapısındaki seri sanal makinenin nöroanatomik yapısı, modüller arasında iletişim kuran uzun aksonlu kapsamlı çalışma alanı nöronlarıyla uyumludur. Kapsamlı nöronal çalışma modelinde paralel olarak her modül işlevini yerine getirirken seçici dikkat veya aşağıdan yukarı dikkat mekanizmaları işlevlerden bazılarını amplifiye, bazılarını inhibe eder. Amplifiye olan aktivasyonlar uzun aksonlu çalışma alanı nöronlarıyla tüm diğer modüllere sunulur. Çalışma alanı nöronlarının aktivasyonuyla tüm modüllere dağıtılan bilgi tutarlı bir şekilde bilinçli deneyimi oluşturur ve bilinçli deneyim dil, uzun süreli bellek oluşumu, istemli eylemler için bir önkoşuldur. Çalışma alanı nöronları sadece kısa bir süre için aktive olabildiği için bilincin içeriği bütünlüklü tek bir deneyimdir. Dennett, ilk olarak kortekste sabit olmayan bağlantı örüntülerinin bir konudan diğerine geçişte hızlıca değişik paketlere geçercesine değişebildiğini belirterek bilinç içeriğindeki bağlantıların hızlıca değiştirilebilir olması gerektiğine dikkat çeker. Dennett, bilinç akışı yerine, çoklu taslaklar modelinde devamlı revize edilen bilinçli deneyim aktivasyonları tanımlamış olduğundan dil, bellek gibi üst seviye modüller, kapsamlı çalışma alanı nöron aktivasyonlarının aktivasyon anlarında söz konusu bilgiye dağıtık bir şekilde bilinçli deneyim içeriklerinin oluşumu esansında ulaşırlar. Aynen belli bir aktivasyonun bilinçli deneyimi oluşturup oluşturmayacağının baştan bilinmeyişi gibi, üst modüllerin söz konusu bilgiyi kullanma olasılığı da aktivasyonun karalılığına bağlıdır. İç içe iki ağ sisteminin birbirinin içine geçmiş olduğunu varsaydığımızda, bu ağ sistemlerinin -tıpkı görsel bilginin oluşumun modüller arasında sabitlenmiş yollara sahip olduğu gibi- birtakım kararlı iletişim yollarına sahip olmasını mümkündür. Örneğin dil sanal makine ağ sisteminin de üç aşağı beş yukarı belli bir düzende değişik modüllerle iç içe geçmiş olmasının ve işlevselliğinin belli bir düzenle açıklanmasının olasıdır. Dolayısıyla beyinde çoklu, üst üstte bindirilmiş işlevsel kümelerin bu şekilde oluşmuş olması mümkündür.

Dennett, kapsamlı nöronal çalışma alanı modelindeki iki seviyeli hiyerarşidense üst seviye bilişsel modüllerin çoklu taslaklar modelinde tarif edilen aktivasyonlarla iç içe geçmiş olduğunu dil kullanımının düşünce süreciyle paralelliğine dikkat çeker. E. M. Forster'ın bir romanından alıntılanmış olan "Söylediğimi anlayana dek ne

düşündüğümü nereden bileyim" ifadesi çoklu taslaklar modelinde üst seviye bilişsel devrelerin beyindeki anlık fark yaratan aktivasyonlarına nasıl ulaştığını betimlemek için doğru bir örnektir. (Dennett, 2011, s: 286)

Dennett, William James'in, evrende bildiğimiz en şaşırtıcı şeyin, bilincin, beynimizin çalışmasında herhangi bir önemli rolü olmayan, sadece yapay bir şey olduğunu varsaymanın saçma olacağı düşüncesine katılır. Genelde bilincin işlevini sorarken bilincin bilgi işleme özelliği üstünde durulması nedeniyle bilincin asıl işlevlerinin gözden kaçırıldığını düşünen Dennett, bilincin evrimde gelecek ön görüsü işlevini yerine getirmek için ortaya çıktığını, çoklu işlevleri birleştirdiğini, ancak bazı özelliklerinin önemli olmadığını belirtir. İnsanlarda bilinç içeriğinin tüm aktivasyonların geçmişe yönelik bir şekilde iz bırakmasıyla ortaya çıktığını hatırlatır. Dennett'a göre bilincin asıl işlevini, kendini izleyerek öz kontrol becerileri geliştirmektir. Olan biteni temsil etme tekniklerimiz, başka hiçbir türün yaklaşamayacağı bir şekilde, bizim öz yöneticiler veya yürütücüler olmamıza izin verir. (Dennett, 2011, s: 325) Geleceğe dair olasılıkları düşünmek, olası senaryoları değerlendirmek uzun vadeli projeler için en gerekli koşullardır. Ayrıca geriye doğru bir bakış, bir hatanın tekrarlanmasını önleyecek en önemli özelliktir. İnsanın kendi kendine konuşma pratiğini geçmişi değerlendirmek için kullanması bir hatadan kaçınmayı, hayvanların öğrenmesiyle kıyaslanmayacak kadar gelişmiş bir seviyeye çıkarmıştır. Dolayısıyla bilincin noktasal bir anda işlevi olmasa bile geniş bir zaman dilimindeki işlevi, organizmanın geçmişini fark ederek gelecek davranışlarını yönlendirmesini sağlamaktır.

Dennett, kendi kendine konuşma pratiğinin bu özelliğinin yüksek dereceden istemli duruşa bir örnek olduğunu belirtir. İstemli duruş, bir sistemin belli rasyonel nedenlere göre hareket ettiğini varsayar. İstemli duruş, bir sistemi incelemek için sistemin her türlü hareketin, eğer sistemin hareketi rasyonel bir çerçevede tahmin edilebiliyorsa, bir nedeni olduğunu varsayar. Ancak bir sistemin davranışlarının nedenleri olması, o sistemin bilinçli bir sistem olduğu anlamına gelmez. Bir bakterinin davranışlarını incelerken, bakterinin belli hareketleri belli nedenlerden ötürü yaptığı düşünülebilir. Satranç şampiyonu Kasparov'u yenen yapay zeka örneği Deep Blue'nun hamlelerinin rasyonel nedenleri bulunmaktadır; Deep Blue oyunu kazanmak için satranç oynamaktadır, ancak Deep Blue'nun bilinçli bir sistem olduğunu düşünmek mantıklı değildir. Bir sistemin bilincinden bahsetmek için o sistemin yüksek dereceden istemli duruş sergilemesi gerekmektedir, ancak inançları hakkında inançları olan bir sistemler, bir başka ifadeyle ikinci dereceden istemli duruş sergileyebilen sistemlerin

bilinçli sistemler olabilirler. Kendi kendine konuşma pratiği inançları dile getirme pratiğidir. Örneğin buzdolabının önünde bekleyen kediyi gördüğünüzde kedinin aç olduğunu söyleyerek kedinin aç olduğu inancını dile getirirsiniz. Ancak ifade edilen bu inanç sadece kedinin aç olduğu inancını içermez, kedinin aç olduğuna inandığınız inancını da içerir. Kedinin aç olması birinci dereceden bir düşüncedir, kedinin aç olduğuna inandığınızı söylediğinizde ise bu ikinci dereceden bir düşüncedir. Ancak ikinci dereceden düşüncenin ifade ettiği birinci derecedeki amacın bilinçli olmasına gerek yoktur. (Dennett, 2011, s: 356) Dennett, Rosenthal'ın, "bilinçli düşüncelerin, bilincinde olunan düşünceler" olduğu görüşüne katılmaktadır.

Dennett, bilincin fenomenal içeriklerini, dolayısıyla *qualia*yı kullanıcı ara yüzü olarak tanımlar. Dennett'ın görüşüne göre, bilincin oluşturulan içerikler bize fenomenal deneyimlerimiz olarak sunulmaktadır, dolayısıyla *qualia*yı daha fazla irdelemenin bir bilinç modeli için faydası olmayacaktır. Bir kullanıcı ara yüzündeki ikonları tasarlamanın birçok yolu vardır. *Qualia* evrimin tasarlamış olduğu ve dış dünyanın gerekli bilgilerini basitleştirilmiş bir biçimde oluşturduğu iç temsiller, yani kullanıcı ara yüzüdür. Bu kullanıcı ara yüzünün, klasik bilgisayarın kullanıcı ara yüzünden tek farkı, ekranda olmayıp, beynin yapısının içinde iş görmesidir. Kullanıcı ara yüzü dış dünyanın bilgilerini sunan bir illüzyondur. Bir illüzyondur, çünkü bilincin içerikleri dış dünya ile ilişkili olsa da sadece beyin tarafından yaratılmıştırlar. Dennett bilincin içeriklerinin oluşmasının bir sihirbazın hilelerine benzetmektedir. İnsan bilinci dış dünyayı oluşturmak için bir çanta dolusu sihirbaz hilesi kullanmaktadır. Bu sihirbaz hileleri, aslında beynin yarattığı mühendislik hileleridir. (Dennett, 2003, s: 18-19)

## 4.4 DENNETT'A GÖRE GÜÇLÜ YAPAY ZEKANIN OLANAKLILIĞI

Dennett güçlü yapay zekanın epistemolojik bir problem olan çerçeve probleminin çözümüyle mümkün olduğunu savunmaktadır. İnsan beyni söz konusu olduğunda çerçeve probleminin çözümü bilinçli olmakla yakından alakalıdır. (Baars, Shananan, 2005, s: 174) Dennett bilince fiziksel duruş pozisyonundan ve işlevsel bir bakış açısıyla yaklaşarak insan bilincinin işlevlerini ve bilincin söz konusu işlevlerinin gözlemlenebilir sonuçlarını incelemiş ve insanların sahip olduğu zihinsel yetilerin ve insan bilincinin işlevlerini yerine getirecek mekanizmaların yapay sistemlerde tasarlanmasının teorik olarak olası olduğunu iddia etmiştir. Ancak her ne kadar teorik olarak insan zihnine eşdeğer özelliklere sahip yapay sistemler yaratmak mümkün olsa

da pratikte bu sistemlerin yaratılması karmaşıklıklarından dolayı mümkün olmayabilir, daha önemlisi ise insan gibi insana eşdeğer yapay sistemler yaratmamıza da gerek yoktur. Dennett hedefin insan benzeri bilinçli sistemler yaratmaktan daha çok işimize yarayacak yapay sistemlerin tasarlanması olması gerektiğini savunmuştur.

Dennett'ın insanların çözmeyi başardıkları çerçeve problemlerini çözebilecek sistemler tasarlamanın teorik olarak mümkün olduğu iddiasının temelinde homunkular işlevselliği bulunmaktadır. Homunkular işlevselliği beynin değişik düzeylerde değişik düzenlilikler gösteren bir yapısı olduğunu savunur. Üst düzey seviyedeki yapıyı meydana getiren bir alt seviyedeki yapı üst düzey seviyenin özelliklerine nispeten sahip ama üst düzey seviyesinde gerçekleşen yetilerden daha azını gerçekleştirme yetisi olan yapılardır. Dennett en üst seviyenin yönelimsel duruşla incelediğimiz zihin seviyesi olduğunu belirtmiştir, zihnin oluştuğu seviye tüm beyindeki alt sistemlerin işlevlerinin toplamıdır. Dennett en alt seviyede ise fiziksel duruşla incelenebilen nöronlar olduğunu belirtmiştir. Dennett, bir nöron sadece belirli bir eşiğin üstünde uyarıldığında ateşlendiğinden dolayı nöronların açık ve kapalı olmak üzere iki duruma sahip makinelerle değiştirilebileceğini, dolayısıyla beynimizdeki nöronları açık ve kapalı durumları olan makinelerle değiştirdiğimizde beynimizin işlevlerini yerine getirmeye devam ederek zihinsel temsillerimizi oluşturacağını düşünür. Ancak bir nöronun beyinde ateşlenmek veya ateşlenmemek dışında başka nöronlarla devamlı değişen bağlanma ağırlıklarının yanı sıra beynin plastik yapısını oluşturacak şekilde yeni bağlantılar kurma özellikleri de bulunmaktadır. Dennett, Bakteriden Bach'a kitabında homunkular işlevselliğini derinleştirerek en alt düzeyde açık ve kapalı durumlarıyla bir makineyle değiştirilebileceğini düşündüğü nöronların özerk nöronlar olduğunu belirtmiştir. (Dennett, 2017, s: 162) Beynimizdeki nöronlar, tüm canlı hücreler gibi içlerinde bulundukları çevrede dengeleşimlerini sağlamaya çalışan özerk hücrelerdir ki nöronların kendi kendini organize eden dinamik yapıları beynin plastikliğini doğrudan açıklamasıdır. Beyinde her bir nöron en az enerji harcayacağı şekilde durmaksızın bağlantıları ayarlamaya çalışmaktadır. Dennett nöronlar özerk hücreler olarak kabul edildiğinde nöronların alt seviyesindeki protein, mikrotübül ve diğer yapıların fiziksel duruşla anlaşılabilir yani işlevlerine göre makinelerle değiştirilebilir olduğunu ve nöronların özerk hücreler olması kabulünün beynin hiyerarşik yapısının ve en nihayetinde en üst seviyede ortay çıkan zihnin açıklamasında bir fark yaratmayacağını savunur. Nöronların özerklikleri daha çok beynin zihni aşağıda yukarı şekilde oluşturuyor olması fikrini vurgulamaktadır. En temel seviyeden zihin seviyesine fiziksel duruşla baktığımızda değişik düzeylerde

yeni özelliklerin belirdiğini belirten Dennett, ancak her düzeydeki yeni özelliğin alt düzeydeki özelliklere üst bağlanır olduğunu düşünür. Dolayısıyla zihni ortaya çıkaran beyin tamamen fiziksel bir yapı olarak taklit edilebilir. İnsan zihnine ev sahipliği yapan beynin organizasyonu aşağıdan yukarı bir tasarımdır ve bu tasarımı tersine mühendislikle yukarıdan aşağı bir tasarımla taklit etmek teorikte mümkündür. Dennett en nihayetinde homunkular işlevselliğe bağlı kalarak insanların robotlardan oluşan robotlar olduğunu belirtir. (Dennett, 2019, s : 110)

Güçlü yapay zekanın olanaklı olmadığını savunan görüşlerden biri zihnin sadece canlı tasarımlarda mümkün olduğudur. Dennett, sadece canlı tasarımların zihni oluşturabileceği görüşünün yanlış bir görüş olduğunu, en nihayetinde DNA'nın bir hücrenin tüm davranışını ve çevresiyle girdiği etkileşimleri tayin ettiğini, fiziksel duruşla organik yapıları yani kendi kendini organize eden yapıların inorganik yapılardan farkının organik yapıların karmaşıklığı olduğunu ve organik yapılara tersine mühendislik uygulamasının oldukça zor ve çetrefilli olduğunu düşünür. Ancak en nihayetinde elimizde fiziksel bir sistem vardır ve bu sistemin işlevi zihnin yerine getirdiği işlevleri yerine getirmektir. (Dennett, 1997, s: 18) Aynı şekilde Dennett bilincin sadece doğan ve gelişen tasarımlara ait bir özellik olmadığını, doğarak dünyaya gelen ve gelişen tasarımların sadece belirli bilgileri taşıdığını ve öğrenme yetilerinin olduğunu, yapay sistemlerde de öğrenmenin mümkün olduğunu ve yapay tasarımlara söz konusu doğuştan gelen bilgilerin el ile konmasının mümkün olduğunu savunur. (Dennett, 1997, s: 19) Güçlü yapay zekanın olanaksızlığını savunan görüşlerin bir kısmı ise insan beyninin yanında yapay sistemlerin çok basit tasarımlar olduğunu belirtirler. Dennett, bir insanın evrim sonucu birçok özellik geliştirmiş milyonlarca hücre topluluğunun bir organizasyonu olduğunu ancak hücrelerden üst düzeydeki belirli sistemlerin işlevlerini yerine getiren tasarımlarla başlayarak en üst seviyeyi tasarlamanın mümkün olduğunu düşünmektedir. Dennett, beyinde en üst seviyenin altındaki seviyelerin tamamen mekanistik olduğunu, dolayısıyla en üst seviyenin altındaki modüllerin tasarımsal duruşla tasarlanmasının mümkün olduğunu düşünmektedir. (Dennett, 2007, s: 78-79)

İnsansı robot Cog projesi Dennett'ın homunkular işlevselliğinin pratiğe dökülmüş bir versiyonu gibidir. Cog insana benzeyen bir robot olarak tasarlanmıştır. Öncelikle Cog doğuştan gelen birçok özellik ile donatılmıştır. Bu özelliklerin bir kısmı Cog'un kendi bütünlüğünü koruması içindir. Ayrıca Cog doğuştan insan yüzlerini tanımaya ve onlara bakmaya programlanmıştır. Cog birden fazla paralel çalışan modüle sahiptir,

bu modüller el göz koordinasyonunu sağlayan modüllerden insan sesini diğer seslerden ayırt eden modüllere çok farklı görevleri yerine getirmek üzere tasarlanmıştır. Ayrıca bu modüllerin bir kısmı Cog'un bazı yetileri kendi başına öğrenmesi için, örneğin dil öğrenmesi için tasarlamış modüllerdir. Dennett Cog'un ona bir eyleme geçmesi için gerekli motivasyonel sistemlerden yoksun olmasının Cog'un en büyük eksiği olduğunu belirtir. (Dennett, 1998, s: 26) Dennett, Cog'un dil öğrenmesi için tasarlanmış, sıfırdan öğrenecek modüllerinin diğer modüllerinde işlevlerini yorumlayarak Cog'un bilincini oluşturacak modüller olduğunu düşünür, teorikte dil öğreniminden sonra Cog'un kendi sanal makinelerini ona verilmiş boş modüllere kurması beklenmektedir.

Dolayısıyla Dennett, Cog'un sisteminde en üst seviyenin altındaki özelliklerin tamamının mekanistik olduğunu ve bu sistemlerin yerine getirecekleri işlevlere göre tasarlanabileceğini, zihnin ise dil sonrası Cog'un sistemine, Cog'un kendisinin kuracağı sanal makineler ile belireceğini düşünmektedir. Dennett'a göre Cog'un dil yetisini atalarımızın binlerce yılda edindiği gibi edinmesi hedeflemiştir. Cog'un doğuştan gelen tasarımlarıyla ve dünyaya dair oluşturduğu iç temsiller ile üst seviyede kendini yeniden tasarlanması umut edilmiştir. Cog başarıya ulaşmamış bir proje olmasa dahi Dennett'ın düşüncelerine göre yapay zeka tasarımının bir örneğidir.

Dennett bilincin iki unsurunu beyin ve kültür olarak tanımlamıştır. Beynin bilgiyi işleyişini çoklu taslaklar modelinde açıklayan Dennett, bilincin birçok özelliğinin kültürle doğmuş olduğunu ve bilincin biyolojik temeli beyin olsa da kültürün bilinci meydana getirmiş olduğunu iddia eder. (Dennett, 1991, s: 35) Beynin biyolojik tasarımı çevreyle etkileşim içinde bütün bir evrim sürecinin sonucu oluşmuştur. Yapay bir sistem yaratmakta ki en büyük zorluk ise milyonlarca yıl içinde evrimin oluşturmuş olduğu bu tasarımın tersine mühendislikle yeniden oluşturulmasıdır. Evrim aşağıdan yukarı bir süreç, bir sistemi tasarlamak ise yukarıdan aşağı bir süreçtir. Evrim kör bir tasarımcıdır, bir başka ifadeyle evrim tamamen fiziksel süreçlerden oluşmakta ve olası tüm tasarımları denenmekte, kötü tasarımlar ise çevreyle girdikleri etkileşimle elenmektedir. Evrimin tasarımlarını mühendislik tasarımlarından ayıran bir özellik ise evrimin her zaman eski sistemlerin üstüne yeni sistemler eklemesidir. Evrim ileriyi hiç görmediği için öngörülemez yan etkileri düşünmez, onları önlemeye çalışmayarak tasarımlarını dener. İki veya daha fazla ilişkisiz işlevsel sistem bir araya gelip, sürprizler üretebilirler. Bu sürprizlerden zeki tasarımcıların tasarımlarında hiç bulunmayan tek ögelerin çoklu işlevleri ortaya çıkar. Çok işlevli tasarımlar aynı

zamandan üst üste eklenen tasarımların da bir yan ürünüdür. (Dennett, 2011, s: 207) Mühendislerse tasarımlarını belli bir amaca yönelik ve aynı zamanda beklenmedik yan etki ve etkileşimleri bertaraf edecek şekilde tasarlar. Ancak Dennett evrimin yarattığı yapıların karmaşıklığına rağmen bu yapıların işlevleri göz önünde bulundurularak tasarımsal duruşla yeniden yaratılmasının teorikte olası olduğunu düşünmektedir. Günümüzde pekiştirmeli öğrenme ile birleştirilmiş evrimsel algoritmalar rastgele yapılarıyla evrim sürecini taklit etmeye çalışmaktadır.

İnsan bilincini yaratan esas unsur beynin fiziksel yapısında kendine yer bulmuş olan insan kültürüdür. Dennett zeka ve idrak yeteneğini birbirinden ayırır. Zeki tasarımlar doğanın her yerinde bulunur, bu tasarımlar evrimin tasarımlarıdır, evrim ise dünyada bilinen en zeki tasarımcıdır. İnsanın dünya hakimi olmasının temeli zekayı idrak etmesidir, insanın zekayı idrak etmesi ise ancak kültürle mümkün olmuştur. İnsanlar, buzul çağının başından sonuna diğer yakın akrabalarına fark atarak beyinlerini dört kat büyütmeyi başarmışlardır. Ancak evrim sürecindeki büyük değişim ve insanların dünyada hakimiyet kurması buzul çağının sonunda insanların dil yetisine sahip olmaya başlamasıyla başlamıştır. Dennett bu büyük değişimin sadece genlerdeki evrimle açıklanamayacağını ve bu gelişimin nedeninin farklı olduğunu söyleyerek değişimin temelinde insan bilincinin ortaya çıkışını yerleştirir. (Dennett, 2011, s: 217) İnsan bilincinin ortaya çıkışı ise insan beyninin düşünce aletleri edinmeye başlamasıyla alakalıdır. Dennett, insan beynin esnek yapısının çevrede karşılaşılan yeni koşullara kendini yeniden organize ederek yanıt vermesini sağlamış olduğunu düşünür. (Dennett, 2011, s: 217) *Homo sapiens* beynini, akrabası olduğu primatların beyinlerinden ayırt eden en belirgin özellik, hacminin akrabalarına kıyasen dört kat daha fazla olmasıdır. Ayrıca *homo sapiens* beynini çevreye göre yeniden organize etmek için diğer primatların sahip olmadığı uzun bir çocukluk dönemine sahiptir. Dennett, ilk zamanlar *homo sapiens*in dil kullanımının diğer primatlarda olduğu gibi değişik seslerle iletişim kurarak başladığını ancak bu dil kullanımının idraktan yoksun bir yetenek olduğunu iddia eder. Zamanla da bazı kelimeler, evcil olmadığı halde insanlarla beraber yaşayan hayvanlar gibi insanların yaşamına girmiş ve insanlar, yabani köpekleri evcilleştirdiği gibi bu kelimeleri de evcilleştirmiş ve ilk düşünce aletleri kullanılmaya başlanmıştır. (Dennett, 2017, s: 197) Bilgi paylaşma işlevi gören dil kullanımı idrak edilmeye başlandıkça kendine soru sorma pratiği, başkalarının sorularını cevaplamanın bir yan etkisi olarak ortaya çıkmıştır. Dil ise *homo sapiensi* günümüze taşıyan kültürün aktarım aracı olmuştur. Kültürün aktarım birimleri memlerdir. (Dennett 2011, s: 415) Memlerde aynen genler gibi evrimleşerek kültürleri

oluşturur, bazı kültürler doğal seçilimle yok olurken bazı kültürler geniş ölçeklere yayılmayı başarır. Dennett, esas olarak bilgiden oluşturulmuş memlerin insan beyinlerine kurulan sanal makineler tarafından işleniyor olmasının akla yatkın olduğunu düşünür. Bilgi soyut bir kavramdır, *bit*lerle ifade edilir, ancak bilgi somut bir yapı aracılığıyla var olmak durumundadır. Her ne kadar çoklu taslaklar modeli bilinç akışından daha çok birbiriyle bağlantılı, geçmiş ve geleceğe göre revize edilen bilinç sahnelerinden oluşsa da, bilinç akışının seri yapısından dolayı memlere ev sahipliği yapacak olan sanal makinenin Von Neuman tarzı seri yapılı bir makine olması gerekmektedir. Beynin paralel yapısında ne hesaplanırsa hesaplansın bilinç beynin hesaplarını aşağı yukarı sıralı bir şekilde takip eder. Sanal makine, beynin fiziksel yapısına bir kez kurulduktan sonra düşünme aletlerini ve sezgi pompalarını içeren yazılımlar taklit, pekiştirme, ödül-ceza gibi yöntemlerle doğumdan başlayarak insan zihinlerine yerleştirilebilirler. (Dennett, 2017, s: 201-202) Basit düşünce aletlerinin üstüne kurulan üst düzey düşünme aletleri, bir konu hakkında nasıl düşünüleceğinin kurallarını içeren sezgi pompaları insan kültürünün en önemli araçlarıdır.

Ayrıca bir sistemin bilincinden bahsetmek için o sistemin yüksek dereceden istemli duruş sergilemesi gerekmektedir, ancak inançları hakkında inançları olan sistemler, bir başka ifadeyle ikinci dereceden istemli duruş sergileyebilen sistemler bilinçli sistemler olabilirler. Dennett'a göre ikinci dereceden düşüncenin ifade ettiği birinci derecedeki amacın bilinçli olmasına gerek yoktur. (Dennett, 2011, s: 356) Birinci dereceki düşünceler ise bilincin oluşturan içerikler, fenomenal deneyimlerimizdir. Dolayısıyla bilinç çoklu taslaklar modeliyle birçok dış ve iç duyusal algının bütünleştirildiği fenomenal deneyim üstüne insanlık tarihi boyunca üst üste kurulmuş sanal makinelerin bir sonucudur. Sanal makineler düşünme aletlerini, sezgi pompalarıyla beynin yapısına gömülü bir şekilde işletmektedirler. Dolayısıyla fenomenal deneyimler sanal makinelerin kullanımına sunulan, organizmanın vücudunun içsel bilgisini ve çevrenin organizmayla alakalı bilgisini basitleştirilmiş bir şekilde içeren, kullanıcı ara yüzleridir.

Dennett bilinci açıklarken bilincin evrim sürecinde ortaya çıkışını açıklamış ve fenomenal deneyimlerin iç ve dış dünyanın bilgisini kullanıcıya sunan bir ara yüz olduğu sonucuna ulaşmıştır. Bilinç ise fenomenal deneyimin içeriklerinin sanal makinelerce işlenmesi sonucu ortaya çıkan işlevsel bir özelliktir. Bilincin ana işlevi ise çerçeve problemini çözmektir. Dolayısıyla bilinci bir mem kompleksi olarak tanımlayan Dennett, güçlü yapay zekanın teorik olarak olası olduğunu düşünmektedir. İnsan

kültürü içinde çerçeve problemini çözebilen bir sistem tanım olarak bilinçli olmak zorundadır. Bu sistemin yaratılması ise ilk adımda evrimle gelişmiş olan doğuştan gelen bilginin tanımlanması, dış dünyadaki bilgilerin ayrı modüllerde tanımlandıktan sonra birleştirilerek kullanılmaya hazır hale getirilmesini kapsar. Bir organizma çevresinde hangi bilgilerin hedefleriyle alakalı, hangi bilgilerin alakasız olduğunu evrim sürecinde gelişmiş olan duygu mekanizmalarıyla algılamaktadır ki pekiştirmeli öğrenme en basit ödül ve ceza olan duygu mekanizmalarının bir taklididir. Dennett bilgiyi kullanılır hale getirenin ise dil olduğunu düşünmektedir. (Dennett, 1984, s: 189) Dolayısıyla Dennett, güçlü yapay zekanın dış dünyanın istatistiki derlemesini yapmakta başarılı olan ve dış dünyanın temsillerini içsel olarak oluşturabilen yapay sinir ağları üstüne seri şekilde kurulacak sanal makinelerle mümkün olduğunu savunmaktadır. Söz konusu sinir ağları beyinde duyusal bilgileri işleyen modüllere denk gelmektedir. Fiziksel tasarıma ilk kurulması gereken sanal makine ise dil anlama ve dili kullanma sanal makinesidir. Dil anlama sanal makinesine sahip olan bir sistem, hedefleri doğrultusunda başka sanal makineleri de kendi başına sistemine kurma yetisine sahip olacaktır.

Dennett'ın güçlü yapay zekayı olanaklı bulması ve bilince bir kullanıcı ara yüzü işlevi vermesinin temelinde ise Dennett'ın yönelimsel duruşu araçsal bir tutum olarak kullanması yatmaktadır. Dennett insan bilincinin açıklanmasına veya güçlü yapay zekanın olanaklılığının araştırılmasına tamamen epistemolojik açıdan yaklaşmaktadır. Dennett bilincin ontolojik özelliklerinin incelenmesinin bilinçle ilgili epistemolojik sorunlar çözülmeden anlamlı bir çaba olmadığını düşünmektedir.

# SONUÇ

Yapay zeka araştırmalarının temeli bilgisayar bilimlerinin öncüsü Alan Turing'in insanların yaptığı her türlü hesabı makinelerin yapabileceğini öne sürmesiyle atılmıştır. Turing bu düşüncesi aynı zamanda felsefede işlevselcilik bakış açısının ortaya çıkarmıştır. İşlevselcilik en temelde zihinsel özelliklerin işlevsel özellikler ve zihnin çoklu gerçekleştirilebilir olduğunu savunur. Dolayısıyla güçlü yapay zekanın olanaklılığından bahsedebilmek işlevselciliğin öne sürdüğü düşünceleri kabul etmekle mümkündür.

İlk işlevselci kuramlardan biri olan makine işlevselliği doğrudan Turing makinelerinden ilham alarak insan zihninin olasılıklı geçişlere sahip bir otomat olduğunu savunmuştur. (Putnam, 1967, s: 54) Rol işlevselliğinin savunucularından Fodor ise temsili zihin kuramında halk psikolojisinin sezgi, inanç, istek, arzu gibi kavramlar üstünden bir model oluşturduğunu belirtmiş, işlevsel bir zihin modelinin üst seviyedeki özellikler üstünden tasarlanması gerektiğini savunmuştur. Fodor'un temsili zihin kuramında düşünce alanını belirleyen nesneler zihinsel temsiller, düşünmek ise bu nesnelere uygulanan önermesel tutumlardır. Fodor, zihnin semantik anlamları olan sembolleri manipüle eden bir Turing makinesine eşdeğer olduğunu iddia etmektedir. Turing makinesinin algoritmasını oluşturan kurallar Fodor'a göre beynin içsel yapısına gömülü kurallardır. (Fodor, 1994, s: 7-9) Fodor'un temsili zihin kuramında zihnin işleyişi alt seviye düzeydeki semantik anlam taşıyan temsillerin bir sembol gibi kullanılarak düşüncenin dilindeki cümleler üzerinden gerçekleşen hesaplamalar ile ifade edilebilir. Önermesel tutumları içeren psikoloji ise beynin yapısından soyutlanmış üst düzey bir bilimdir. Fodor'un işlevselliği sembolik sistemlerin temelindeki işlevselliktir.

Dennett arzu, inanç gibi zihinsel durumları organizmalara veya yapay sistemlere atfederek birçok organizma ya da yapay sistemin davranışlarını bu kavramlar üstünden öngörebileceğimizi ancak sistemlere yönelimsel duruşu atfeden biz olduğunu vurgular. Bir sistemin davranışları yönelimsel duruşla açıklandığında ya da yönelimsel duruşla öngörüldüğünde bu sistem yönelimsel duruşu olan bir sistemdir. (Dennett, 1987, s: 2-3) Dennett, aynen bilgisayara yönelimsel duruş atfettiğimiz gibi hayvanlara ve insanlara da yönelimsel duruş atfetmenin onların davranışlarını açıklamak için iyi bir yöntem olduğunu savunmuştur. Ancak Dennett, bir sistemin

gerçekten yönelimsel bir duruş sergilemesiyle bir sistemin davranışının yönelimsel bir duruş atfedilerek öngörülmesi arasında fark olduğunu da belirtir. İnsanların yönelimsel duruşa sahip olduğunu ve insanların rasyonel ajanlar olduğunu varsayarız. İnanç ve arzularla tanımlanabiliyor olmak her zaman inanç ve arzulara sahip olmak değildir ve bir sistemin gerçekten inanç ve arzulara sahip olup olmadığını anlamamızı sağlayacak bilimsel bir yöntem yoktur. (Dennett, 1987, s: 8) Dennett, birinci dereceden yönelimsel duruşa sahip olmayı sistemin davranışının inanç ve arzular üstünden öngörülebilir olmasıyla tanımlar. İkinci dereceden yönelimsel bir duruşa sahip olmaksa sistemin kendi inanç ve arzularına dair inançları olmasıyla tanımlanır. Dennett gerçek anlamda yönelimsel duruşa sahip olan sistemlerin ikinci dereceden yönelimsel duruşa sahip olan sistemler olduğunu belirtir. Bu şekilde Dennett insan zihni ile yaratık zihni arasında bir ayrım yaparak yaratık zihninin yaratığı yönlendiren temsillerinin açıklaması için tasarımsal duruşun yeterli olduğunu savunur. Dolayısıyla Dennett zihinsel kavramları sadece sistemlerin davranışlarının açıklanması veya öngörülesi için kullanışlı, araçsal kavramlar olduğunu savunmaktadır.

Bilim ise fiziksel duruşla sistemlerin davranışlarını matematiksel yöntemlerle öngörürler. Her bilim dalının kendi alanlarının olmasının nedeniyse sistemlerin karmaşıklık dereceleri arttıkça beliren yeni fenomenlerdir. Yeni fenomenlerin hareketlerinin öngörülmesi için her bilim dalı kendi kavramlarını kullanmaktadır. Örneğin temel kuantum fiziği atomların oluşumlarını ve hareketlerini kendi kavramlarıyla öngörürken kimya bilimi atomların birbiriyle etkileşimlerini açıklamak için başka kavramlar kullanmaktadır. Bunun nedeni kimya bilimini fiziğin kavramlarıyla açıklamaktaki karmaşıklıktan kaynaklanan zorluklar ve kimya biliminde fizik kavramlarının kullanılmasının verimli bir yöntem olmayışıdır. Aynı şekilde biyoloji canlıları kendi kavramlarıyla açıklar ve fiziksel duruş yerine tasarımsal duruşu benimser ancak DNA kimya bilimiyle açıklanabilen bir fenomendir.

Dennett psikoloji biliminin üst seviye bir bilim olduğunu ve kendi kavramlarıyla iş gördüğünü belirtir. Psikoloji biliminin inanç, arzu gibi kavramları belirli düzenlilikleri de içermektedir. Ancak bu kavramların gerçekliğini sorgulamak çok anlamlı bir çaba değildir. Beyin evrendeki en karmaşık sistemdir ve Dennett zihnin gerçek yapısını anlamak için beynin farklı düzeylerden oluştuğunu ve her düzeyde farklı düzenlilikler sergilediğini kabul etmemiz gerektiğini düşünür. Sistem içindeki bir fenomen sistemin içinde, bütün bir sistemin yapısı hesaba katılarak incelendiğinde, hesaplanamaz olsa da sistemin tüm özellikleri fiziksel duruş tarafından teorikte öngörülebilmektedir.

Ancak bütün bir sistemi parçalayıp, birbiriyle etkileşen sistemlerden birini tek başına incelediğimizde, özellikle kendini organize eden sistemlerde, sistemin çevresiyle etkileşiminden dolayı sanki sistem yukarıdan aşağı bir nedensellik gösteriyormuş gibi gözükür. Bu durum özellikle biyoloji ve zihin için geçerlidir. Zihne atfettiğimiz inanç, arzu gibi kavramlar da zihni çevreden bağımsız bir fenomenmiş gibi, yönelimsel duruştan incelememizden dolayı bize olduklarından farklı gözükebilirler. Ancak bu kavramların araçsal kavramlar olduğunu unutmamalı ve bu kavramlarla ifade edilen fenomenlerin fiziksel olmayan, farklı ontolojilere sahip olduğunu düşünmemeliyiz.

Dennett zihne homunkular işlevselliği ile yaklaşır. Homunkular işlevselliği zeki bir sistemin, yani zihnin, sistemin bütününden daha az zeki birçok alt sistemin birlikte işlev görmesiyle açıklanabileceğini savunur. Dennett en temel seviyede makineler ile değiştirilebilecek açık veya kapalı durumda olan nöronların en üst seviyede zihni ortaya çıkardığını savunmaktadır. Ayrıca Dennett'a göre zihnin belirdiği en üst seviye dışındaki alt seviyeler tamamen mekaniktir ve bu düzeyler tasarımsal duruşla incelenmeye yani işlevlerini yerine getiren tasarımlar olarak kabul edilmeye uygundurlar. Dennett'ın özellikle zihnin kavramlarının araçsal kavramlar olduğunu savunması ve bu kavramlara sadece öngörü kapasitelerine göre önem vermesi, bir başka ifadeyle zihinle ilgili çözülmesi gereken tüm problemin epistemolojik problem olduğunu savunması güçlü yapay zekanın olanaklılığını epistemolojik bir probleme indirgemiştir ki Dennett güçlü yapay sistemler tasarlamanın mümkün olduğunu düşünmektedir.

Başlangıç tarihinden itibaren yapay zeka çalışmalarının da hedefi insan zihninin özelliklerine sahip makineler yapmak olmuştur. Ancak yapay zeka araştırmaları başlangıç tarihinden itibaren iki odağa ayrılmıştır. Bağlantıcılığı savunan yapay zeka araştırmacıları, insan beyninin nöronlar ve nöron devreleriyle iş gördüğünü dolayısıyla nöron devrelerini taklit eden mimariler ile insan zihninin yetilerinin taklit edilebileceğini düşünmüşlerdir. Sembolik sistemleri savunan yapay zeka araştırmacıları ise insan zihninin kurallı işleyişine dikkat çekmiş, özellikle insan dilinin kurallı yapısının insan zihninin kurallarla işlediğinin ispatı olduğunu öne sürmüşlerdir. 1980'li yıllardan itibaren Chomsky'nin dilsel yeterlilik kuramına karşı çıkan görüşlerin yanı sıra nörobilim araştırmalarındaki gelişmelerin de etkisiyle yapay zeka çalışmalarının odağı yapay sinir ağlarına kaymıştır. Günümüzde yapay zeka çalışmalarının odağı büyük ölçüde yapay sinir ağı çalışmalarına kaymış olsa da yapay sinir ağlarının ciddi eksiklikleri bulunmaktadır. Üretkenlik, sistematiklik, birleşimsellik ve çıkarımsal

tutarlılık yapay sinir ağlarının henüz sahip olmadığı özelliklerdir ki bu özellikler insan bilişinin genelleme yaparak öğrenme, çok kısıtlı sayıda öğrenme, bilinmeyen bir bilgiyi, benzer bilgilerden çıkarımsama gibi yetilerinin temelini oluşturur. (Fodor, Pylyshyn, 1988) Yapay sinir ağları her ne kadar istatistiki kuralları alt-sembolik seviyede oluşturma yetisine sahip olsa da  zihnin dört temel özelliğine sahip olmadığından kuralları genelleyememektedir. Ancak çoğu yapay zeka araştırmacısı bu sorunların beynin alt sistemlerinin yapay sinir ağlarına yeni mimariler ile entegre edilmesi ile çözüleceğini düşünmektedir. Örneğin Wang tekrarlayan sinir ağlarını pekiştirmeli öğrenme ile birleştirerek sistemin aktivasyon desenlerini kullanarak basit soyut kuralları öğrenmesini sağlamıştır. (Wang, 2018) Wang'ın sistemi tekrarlayan sinir ağları ve pekiştirmeli öğrenmeyi birleştirerek insan beynindeki çalışan belleği tekrarlayan sinir ağlarında oluşturmayı başarmıştır. Omurgalılar birden fazla bellek sistemine sahiptir. Omurgalılarda epizodik bellek hipokampüs tarafından oluşturulur. Hipokampal indeks modeline göre kortekste oluşan her desen hipokampüste, Hopfield ağlarını andıran paralel devreler ile ilişkilendirilir. Desen tamamlama denen bu işlem sonrası hipokampüste bilinçli deneyimin yaşandığı anda kortekste oluşan desenin benzersiz bir haritası oluşturulmuş olur. Hipokampüsün devrelerinde anıya ait içerikler saklanmaz, içerikler kortekse yayılmış nöronal haritalarda saklanır, hipokampüs sadece bu haritaların o anı esnasında nasıl ilişkilendiklerinin indeksini tutar. (Teyler, Rudy, 2007, s: 6) Derin pekiştirmeli ağlara sembolik bir epizodik bellek eklenmesinin de ağlara tek seferde öğrenme gibi özellik kattığı gösterilmiştir. (Lake, 2017) Ayrıca beyinde soyut haritalar oluşturan ızgara hücrelerinin bulunması ki ızgara hücrelerinin beyindeki soyut düşünmeyle sıkı bir ilişki içinde olduğu düşünülmektedir ve bu durum beynin birçok üst seviye fonksiyonunda basit nöron ağlarından daha karmaşık mimariler kullandığının bir ispatıdır. (Hawkins, 2019, s: 11)

Yapay sinir ağı çalışmaları çoğunlukla nörobilim araştırmalarından ilham alarak değişik ağ mimarileri oluşturmaktadır. Örneğin Bayesci modellerden ilham alınarak tasarlanan tekrarlayan sinir ağlarının beynin çalışma prensiplerini taklit ettiği düşünülmektedir. Friston, Bayesci modellerin nasıl iş gördüğünü "serbest enerji prensibi" ile fiziksel bir temele oturtmuştur. Kendi kendini organize eden tüm biyolojik yapıların kendi bütünlüklerini korumak için dış dünya ile dış dünya dair kendi iç temsilleri arasındaki ön görü farkını en aza indirmeye çalışır gibi bir davranış sergilediklerini iddia eder. Beyin duyusal algıdan doğru çıkarımlar yapabilmek için ilk olarak iç temsillerinde dış dünyanın bir modelini üretmek zorundadır. Beynin dış dünyayı kendi iç temsillerinde ürettiğinde, her yeni duyusal algıyı önceden oluşturmuş

olduğu iç temsilleriyle kıyaslayarak, iç temsillerinin dış dünyayı doğru ürettiğine dair kanıtlar aramaktadır. Fiziksel olarak bu durum beynin dış duyusal bilgi ile iç temsilleri arasındaki farkı, yani serbest enerjiyi minimize etmeye çalışması demektir. (Friston, 2012, s: 1232-1233)

Tekrarlayan sinir ağlarının en ilgi çekici özellikleriyse dış dünyanın temsillerini içsel modellerinde yaratmalardır ki temsilciliğe göre bir şeyi bilinçli olarak deneyimlememiz, o şeyin zihinsel temsili ile ilgilidir. Tüm zihinsel temsiller bilinçli değildir ancak bilinçli her deneyim beynin içsel temsillerini içerir. Eğer ki beynimizde yukarıdan aşağı bir işleme söz konusu ise, yukarıdan aşağı işlemeyi meydana getiren zihinsel temsilin bilinçli bir içerik oluşturması muhtemeldir. Bilinçli içerikler her zaman için ikinci dereceden temsillerdir; dış dünyanın yaratılmış içsel temsilinin tekrardan temsilleridir.

Her ne kadar yapay zeka araştırmacıları bilinçli makineler tasarlamanın mümkün olduğunu düşünse de Searle yapay sistemlerin çok karmaşık hesaplar yapabilen makinelerden öteye geçemeyeceklerini iddia etmiştir. Searle Çince odası düşünce deneyi ile yapay sistemlerin sadece sentaksa sahip olabileceğini ancak asla semantiğe sahip olamayacağını ispatlamıştır. Searle'ün Çince odası düşünce deneyi sembolik sistemler düşünülerek tasarlanmıştır ve temelde Searle'ün bahsettiği sorun sembolik sistemlerdeki sembol temellendirme sorunudur.

Dreyfus ise fenomenolojik bir özne olmanın anlamından yola çıkarak güçlü yapay zekanın mümkün olmadığını, insan zihninin dünyayı anlamlandırmasının mantık kurallarıyla gerçekleşmediğini, insanların ancak dünyanın içinde var olarak dünyayı anlamlandırdığını, dolayısıyla belli kurallarla bir zihin yaratılamayacağını belirterek fenomenolojik öznenin zihinsel temsillerinin çeşitliliğine dikkat çekmiştir. Dreyfus, insan zihninin yapısı bir şekilde biçimsel formüllere indirgense dahi insanın yaşam boyu kazandığı becerilerin sayısal verilere indirgenmesinin imkansız olacağını savunur. (Dreyfus, 1972) Dennett, Dreyfus'un belirttiği problemin, açıkça tanımlanmamış olsa da "çerçeve problemi" olduğunu belirtir. (Dennett, 1984, s: 184) İçsel temsiline göre dünyayı ve dünyadaki ilişkileri tanımlayan bir program, dış dünyadaki durumsal değişikliklerde iç temsillerinde ve iç temsillerindeki ilişkilerde doğru bir güncelleme ile başa çıkmak durumundadır. Çerçeve problemi bir sistem çevre ile etkileşime girdiğinde sistemin içsel aksiyomlarından hangilerinin değişip hangilerinin sabit tutulması gerektiğinin çözümlenmesidir. Dennett güçlü yapay zekanın olanaklılığının derin bir epistemolojik problem olan çerçeve probleminin çözümünün olanaklılığıyla doğrudan ilişkilendirmiştir. Dennett çerçeve problemini

parçalara ayırmıştır. Çerçeve problemi ilk olarak belirli bir durumda hangi bilgilerin o durumla alakalı, hangi bilgilerin durumla alakasız olduğunun tayin edilmesidir. İkinci adımda durumla ilgili bilgileri belirleyen bir ajanın hareketlerinin sonuçlarını tahmin etme yeteneği olması, hareketin sonucunda dış dünyanın temsillerinin iç temsillerini nasıl değiştireceğini tahmin edebilmesi gerekir. Bir başka sorunsa dünyanın robotun hareketlerinden de bağımsız olarak devamlı değişmesidir. Aynı zamanda çerçeve problemi bir zeki ajanın neleri bilmesi gerektiğini ve bu bilgileri nasıl elde edeceğiyle de ilgilidir ki bu adımda biyolojik bir organizma olarak sahip olduğumuz iki özellik devreye girer, ilki doğuştan sahip olduğumuz bilgiler, DNA'mızla belirlenmiş beyin yapımız. İnsanın öğrenmesinin bir kısmı gözetimsiz bir öğrenme ile temel fizik yasalarını, nedensellik gibi kavramları öğrenmesini kapsar ama öğrendiğimiz bilgiler çoğunlukla dil aracılığıyla içinde bulunduğumuz kültür aktarımıyla ilgilidir. İnsanlar bir kültürün içinde doğar ve o kültürün sahip olduğu düşünce aletlerini çoğunlukla dil aracılığıyla öğrenirler. Dennett bilginin ifade edilmesinin ve kullanılabilir hale getirilmesinin dil ile gerçekleştiğini ifade eder. (Dennett, 1984, s: 189)

Günümüzde tekrarlayan sinir ağlarıyla birleştirilmiş derin pekiştirmeli öğrenme, kapalı sistemlerde bir ajanın hedefleri doğrultusunda çevrelerindeki düzenlilikleri öğrenmesini ve ön görmesini başarabilmektedir. Çerçeve problemi açısından bu ajanların en önemli eksiklikleri genelleme yetenekleridir ki bunun sonucu bu akanlar kapalı sistemlerde gösterdikleri başarıyı açık sistemlerde gösterememektedir. Günümüzde çoğu yapay sistem hayatına bir modelden bağımsız, yani doğuştan hiçbir bilgiye sahip olmaksızın başlamaktadır. Ayrıca dile sahip olmayan ve dil ile öğrenmeyen herhangi bir sistemin insan dünyasında başarılı olma sansı olmayacaktır. Günümüzde yaşadığımız dünyayı hatta kim olduğumuzu belirleyen kültürümüzde bulunan ögelerdir. Dolayısıyla çerçeve problemini çözen bir ajan tanım olarak bilince sahip olacaktır.

Bilincin bilim tarafından çalışılabilir olduğunu savunan Crick, bilincin beynin bir ürünü olduğu belirterek, bilincin bilimsel metotla çalışılabilir olduğunu öne sürmüş ve bilinç araştırmalarının temelinin bilincin nöronal korelasyonlarını bularak oluşturulması gerektiğini savunmuştur. Günümüzde kapsamlı çalışma alanı modeli en yaygın kabul gören bilinç modellerinden biridir. Kapsamlı çalışma alanı modeli beyindeki bilişsel süreçlerin çoğunun bilince gerek duymaksızın gerçekleştiğini, ancak belli bir süre belli kararlılığa sahip aktivasyonların bilinçli içerik oluşturduğunu ve seçici dikkatin bilinç için ön koşul olduğunu belirtir. Beyinde belli bir süre belli kararlı durumda kalan

aktivasyonlar seçici dikkat mekanizmalarıyla amplifiye oldukları taktirde bu aktivasyonların taşıdığı bilgi uzun aksonlu çalışma alanı nöronlarınca tüm beyne dağıtılır. Çalışma alanı nöronlarının çift yönlü aksonları aynı zamanda tüm beyne dağıtılmış bilgiyi kullanarak alt modülleri aktive eder ve kararlı aktivasyonların birbiriyle etkileştiği paralel işlemeler seri bir etkinlik oluştururlar. Kapsamlı çalışma alanı modeline göre bilgi global olarak kullanılabilir şekilde korteksin değişik modüllerine dağıtıldığında, bu bilgi aynı zamanda öznel bilinçli deneyimi yaratmaktadır. (Deheane, 2018, s: 218) Bir başka bilinç modeli olan entegre bilgi teorisi ise fenomenolojik deneyimin özelliklerini analiz ederek deneyimin beyinde hangi mekanizmalarla yaratıldığını araştırmıştır. Bilinçli deneyimin tüm sistemde fark yaratan farklılaşmalar bir başka ifadeyle bilinçli içeriğin oluşması için gerekli koşulun beyindeki aktivasyonun daha küçük parçalarla tanımlanamayan ve tüm beyne yayılmış, bilgi içerikli aktivasyonlar olması gerektiğini öne sürmüştür. Entegre bilgi teorisine göre tekrarlayan sinir ağlarının bilince sahip olması muhtemeldir. Nörobilim araştırmacıları elde ettikleri başarılarla ileride bilinçli deneyimin özelliklerini açıklığa kavuşturacaklarını düşünmektedir.

Chalmers nörobilim araştırmacılarının uğraşmakta oldukları problemin kolay problemler olduğunu, bilincin anlaşılması için çözüm bekleyen problemin ise bilincin zor problemi olduğunda ısrarcıdır. Bilincin zor probleminin neden öznel deneyimlerimiz olduğunu, öznel bilinç içeriklerimizin fiziksel bir sistem tarafından nasıl yaratıldığı sorularıyla tanımlar. (Chalmers, 1995, s: 202) Bilincin zor problemi yani fenomenal deneyimlerimizin nasıl oluştuğu felsefe dünyasını ikiye bölmüştür. Bazı filozoflar bilincin fiziksel olmayan farklı bir töz olduğunu savunurken bazı filozoflar bilincin, beyin karmaşıklığından dolayı belirdiğini ve bilincin fenomenal yönünün indirgemeci yöntemlerle açıklığa kavuşturulacağını savunmuşturlar. Nagel ise bilincin birinci sahsın öznel deneyimi olduğunu ve eğer bilim bilincin nöral korelasyonlarını bulsa bile bilincin birinci sahsa verilişinin üçüncü şahıslar tarafından anlaşılamayacağını belirtmiştir. Dennett indirgemeci bir filozof olarak bilinç problemine epistemolojik olarak yaklaşmaktadır. Dennett'a göre bilincin zor problemini yaratan durum bilinç problemine yaklaşım biçimimizdir. Dennett *qualia* kavramının varlığını sorgulamaz ancak bu kavramın fenomeloglar tarafından yanlış anlaşıldığını, *qualia'*nın temel bir kavram olmadığını düşünmektedir. *Qualia* kavramı evrendeki diğer temel kavramlar gibi tanımlanmış kendine has değişmeyen özellikleri olan, başka kavramlarla açıklanamayan bir kavram değildir. Eğer *qualia*yı zihnin ontolojisine dahil edeceksek, qualia'nın kararlı ve tanımlanabilen özellikleri olması

gerekir. (Dennett, 1988) Dolayısıyla Dennett bir bilinç modelinde *qualia'nın* açıklanması gerekmediğini, çünkü *qualia'nın* temel bir kavram olmadığını savunur..

Dennett, çoklu taslaklar modeliyle bilinci açıkladığını iddia etmiştir. Dennett'ın çoklu taslaklar modeli kapsamlı çalışma alanı modelini temel alır ve kendi bilinç modelini çoklu taslaklar modeli üstüne kültürün ögelerini ekleyerek tanımlar. Dennett, *qualia* içeriklerinin zihnin temel bir özelliği olmadığını, dış dünyada var olan, organizma için önemli bilgileri organizmayla ilişkilendiren nitelikler olduğunu ve bu niteliklerin birinci şahsa ne şekilde verildiğinin bir öneminin bulunmadığını iddia eder. Dennett'e göre, beynin *qualia*yı nasıl oluşturduğunun anlaşılması yeterlidir, *qualia* evrimin tasarladığı kullanıcı yüzüdür. Bilinci mümkün kılan dil ile edinilmiş düşünce aletleridir. Dil yeteneği sayesinde insanlık birçok düşünce aleti geliştirmiş, en önemlisi kendisini içinde bulunduğu kültür ve söz konusu kültürün düşünme aletleri aracılığıyla tanımlayarak kendini idrak etme yeteneği kazanmıştır. Dennett bilincin saf bir beyin üstüne yığılmış bir mem yığınından oluştuğunu iddia etmektedir ki mem yığınları bilgi ve bilgiyi kullanma kurallarını içerir. Dolayısıyla Dennett bilinci mümkün kılanın beynin biyolojik yapısı üstüne kurulmuş bilgi ve bilgi kullanım kuralları olduğunu belirtmiş ve sanal makinelerin bilgi kullanım kurallarını içeren makineler olmasından dolayı bilincin beynin biyolojik yapısına gömülü sanal makineler ile oluştuğu sonucuna varmıştır.

Bilinci kültürle tanımlayan Dennett, içsel temsillerini dinamik olarak yaratan ve güncelleyen, alaka problemini hedefleri doğrultusunda çözebilen tasarımların ancak sanal makinelere sahip olduklarında çerçeve problemini çözebileceğini belirtir. Bilinci yaratan insan beyni paralel donanım üstüne gömülmüş seri sanal makinelerden oluşur dolayısıyla kendi kontrol sistemlerini yaratacak paralel donanıma sahip bir makinenin paralel donanımı üstüne kurulmuş sanal makinelere sahip herhangi bir sistem bilinçli olabilir. Bilincin ikinci dereceden bir düşünce ve *qualia*nın bir illüzyon olduğu kabul edilirse ikinci dereceden düşünceler üretebilen bir sistem, kendisini tanımlayacak bir illüzyon yaratma yetisine sahip olacaktır. Ancak böyle bir sistemin yaratacağı illüzyon, muhtemelen, organik sistemlerin evrimle şekillenmiş renklerden, seslerden, duygulardan, hislerden oluşan *qualia* illüzyonundan çok farklı olacaktır.

Dennett bilincin dilden önce beliren ve dili mümkün kılan bir özellik olmaktan daha çok dil kullanımıyla beliren bir özellik olduğunu savunmaktadır. Bir açıdan hetero-fenomenolojik yöntemde ancak dil yetisine sahip ajanlara uygulanabilen bir yöntemdir. Kurallara sahip bir dil kullanımının canlılar arasında sadece insana ait bir yeti olduğu açıktır. Ancak üst düzey memelilerin benlik bilincine sahip olması, deniz

memelerinin ve primat topluluklarının kültürleri olması Dennett'ın bilinci üst düzey bir dil yetisiyle bağdaştırmasıyla açıklanamamaktadır. Dennett, benlik bilincinin dahi kültürle oluştuğunu öne sürer; "Biz neyiz sorusunun yanıtı, kültürün bizi ne şekilde etkilediğinde yatar." (Dennett, 2011, s: 406) Bunun nedeni ise Dennett'ın gerçekten yönelimsel duruşa sahip sistemlerin ikinci dereceden yönelimsel duruşa sahip sistemler olarak kabul etmesinden kaynaklanır. Ancak Dennett'tan farklı bir yol izleyerek önce bilinci tanımlayıp, bilincin işleviyle dil kullanma yetisi kazanan sistemlerin gelişmiş düşünce aletleri edinerek çevreyle girdikleri ilişkide lineer olmayan yeni bir düzenlilik sergilemeye başladığını ve bu yeni düzenliliğin ise McGreedy patlamasını açıkladığını iddia edilebilirdik. Ancak Dennett'ın zihin felsefesi dili bilinç üstüne kurulan bir yapı kabul ederse, hetero-fenomenolojik yöntemi belli sistemlere uygulanamayan eksik bir yöntem olduğunu kabul etmek durumundadır. Dennett bilincin çalışma yönteminin hetero-fenomenolojik yöntem olduğunu savunmuştur. Hetero-fenomenolojik yöntem birinci şahsa sunulan bilinç içeriklerin üçüncü şahıslar tarafından çalışılabilir kılar. Dolayısıyla Dennett'ın hetero-fenomenolojik yöntemi benimsemesi Dennett'ın dili bilincin önkoşulu kabul etmeye mahkum bırakır.

Dennett bilinci temelde homunkular işlevselliği kabul ederek incelemiş ve zihnin beynin bütününde yani en üst seviyede beliren bir özellik olduğunu savunmuştur. Bu özelliğin belirimini ise beynin biyolojik yapısına kurulan sanal makinelere bağlamıştır. Sanal makineler tanımlanmış yazılımlardır ve sistemdeki bilginin tanımlanan kurallar aracılığıyla işlenmesine olanak tanırlar. Dennett kapsamlı çalışma alanı modelinde beynin hesaplamalarını birleştiren uzun aksonlu nöronların söz konusu sanal makinelerin kurulumu için uygun bir yapı oluşturduğunu, düşüncenin seri yapısını ise bu fikrini desteklediğini belirtmiştir. Ancak kapsamlı çalışma alanı nöronları memeli beyinlerde bulunan bir yapıdır ve dil olmaksızın da bu nöronlara sahip beyinler edindikleri içsel ve dışsal duyusal bilgileri seri şekilde işlemektedirler. Dil yetisi bilgi işlenmesi esnasında devreye girmekte ve işlenen bilginin ifade edilmesini sağlayarak yeni bir düzenlilik oluşturmaktadır. Dil yetisinin bir başka etkisi ise zihnin bilgi işleme kapasitesini arttıran düşünce aletlerinin edinilmesini sağlamaktır. Dil *homo sapiensi* diğer canlılardan ayırt eden düşünce aletlerine, hatta nasıl düşünüleceğine dair ikinci dereceden düşünme aletlerine sahip olmasını sağlamıştır. Düşünce aletlerine ve nasıl düşünüleceğine dair ikinci dereceden düşünce aletlerine sahip bir beynin ise bilgi işleme kapasitesinin üstel bir fonksiyonla arttıracağı kesindir.

Dennett, homunkular işlevselliğini kabul etmiş ve zihni en üst düzeyde beliren bir fenomen olarak kabul etmiş ve sadece en üst düzeyde sisteme atfedilen yönelimsel duruşun sistemin gerçekten sahip olduğu bir duruş olduğunu belirtmiştir. Ancak dilin bilinç için ön koşul değil de dilin bilinci olan bir sistemin edinebildiği bir araç olduğu kabul edilirse yönelimsel duruş beynin en üst seviyesinin bir alt seviyesinde beliren bir duruş olduğunu kabul etmek gerekir. Bu kabulde ise entegre bilgi teorisinin bilinci bilgi ile eşleştirmesi bilgi kavramının ontolojik temelleri sorgulanmadan bir araç olarak kabul edilebilir ve böylece dil öncesi gelişen nedensellik gibi bilinçli kavramların oluşumunun incelenmesi mümkün kılınır. Dennett yönelimsel duruşu sadece ikinci dereceden yönelimsel duruşa sahip sistemlerin gerçekten sahip olduğu bir duruş olarak kabul ederek aç gözlü bir indirgemecilik yapıyor izlenimi vermektedir. Yönelimsel duruşu zihnin en üst seviyesinden bir alt seviyeye atfetmek çerçeve probleminin çözümünde bir fark yaratmayacaktır. Yönelimsel duruşun hangi seviye de gerçek bir yönelimsel duruş olduğuna dair geliştirilecek bir modelin yapay zeka çalışmalarındaki önemi ise hayvan haklarına verilen önemle aynı seviyede olacaktır. En nihayetinde bilinç herhangi bir şeyin neden önemli olabileceğine karar veren hakemdir.

# KAYNAKÇA

Baars. In the Theatre of Consiousness, Oxford University Press. 1997.

Baars., Shanahan. Appliying Global Workspace Theory to the Frame Problem. Cognition 98 (2): 157-176. 2005.

Bahdanan. D., Cho. K., Bengio. Y. Neural Machine Translation by Jointly Learning to Alingn and Translate. Conference paper ICLR. 2015.

Bellman. The Theory of Dynamic Programming. Bulletin of the American Mathematical Society, no: 60, pp: 503-515. 1954.

Bengio. Y., Lamblin, P., Popovici. D., Larochelle. H. Greedy Layer-Wise Training of Deep Networks. Conference Paper in Advences in Neural İnformation Processing System. 2007.

Brooks, Rodney. Elephants Dont Play Chess. Robotics and Autonomous Systems 6. 1990.

Caramaza, A. On drawing inferences aboıt the structure of normal cognitive systems from the analysis of patterns of impaired performance: The case for single-patient studies. Brain and Cognition, 5(1), 42-66. 1986.

Chalmers, David. Facing Up to the Problem of Consciousness The Journal of Consciousness Studies 4(1):3-46. 1995.

Chalmers, David. Consciousness and its Place in Nature. 2003.

Chomsky, Noam. Aspects of the Theorty of Syntax. MIT Press. 1965.

Churchland, Patricia. Nörofelsefe. Alfa basım. 2019.

O'Connor, Timothy, "Emergent Properties", The Stanford Encyclopedia of Philosophy. 2020.

Crick, F. C. Şaşırtan Varsayım. Tübitak Yayınları. 1997.

Crick, Koch. Towards a neurobiological theory of consciousness. Seminars in the neurosciences, Vol 2, 263-275. 1990.

Crick, Koch. Are we aware of neural activity in primary visual cortex? Nature , 375, 121–123. 1995.

Damasio, Antonio. Spinoza'yı ararken. ODTÜ Geliştirme Vafkı Yayıncılık. 2018.

Damasio, Antonio. Zihindeki benlik. ODTÜ Geliştirme Vafkı Yayıncılık. 2020.

Deheane, S., Naccehce. L. Towards a cognitive neuroscience of consciousness: basic evidence and a workspace framework. Cognition 79. 2001.

Deheane, Stanislas. Bilinç ve Beyin. Alfa Basım. 2018.

Deheane, S., Naccehce. L. Towards a cognitive neuroscience of consciousness: basic evidence and a workspace framework. Cognition 79. 2001.

Dennett, D. On the absence of phenomenology. Body, Mind, and Method. pp: 93-113. 1979.

Dennett, D. Cognitive Wheels: The Frame Problem of AI, 1984.

Dennett, D. The Logical Geography of Computational Approaches: a View from the East Pole. in R. Harnish and M. Brand, eds, The Representation of Knowledge and Belief, Tucson: University of Arizona Press. pp: 59-79. 1986.

Dennett, D. Intentional Sentance. 1987.

Dennett, D. Quining Qualia. A. Marcel and E. Bisiach, eds, Consciousness in Modern Science, Oxford University Press, 1988.

Dennett, D. Darwin'in tehlikeli fikri, Alfa basım, 1995.

Dennett, D. Consciousness in human and robot minds. in Masao Ito, eds, Cognition, Computation, and Consciousness. Oxford University Press. 17-29. 1997.

Dennett, D. "Are we explaining consciousness yet?". Cognition 79. 221-237. 2001.

Dennett, D. Explaining the "Magic" of Consciousness. Journal of Cultural and Evolutionary Psychology, 7-19. 2003.

Dennett, D. Real Patterns. in Humpereys eds Emergence. MIT Press. 2008.

Dennett, D. The Cultural Evolution of Words and Other Thinking Tools. Cold Spring Harb Symp Quant Biol published online August 17, 2009.

Dennett, D. Bilinç Açıklanıyor, Alfa Basım, 2011.

Dennett, Daniel From Bacteria to Bach and Back: The Evolution of Minds. W.W. Norton & Co. 2017.

Dennett, D. *A* History of Qualia, Topoi, October 23rd, pp. 1-8. 2017.

Dennett, D. "What can we do?" in J.Brockman ed., Possible Minds: 25 Ways of Looking AI. Chapter 5. 2019.

Dreyfus, Hurbert. What Computers Can't Do. Harper&Row Pub. 1972.

Dreyfus, Hurbert, Dreyfus, Stuart. Mind over Machine. The Free Press. A Division of

Macmillan, Inc. Newyork. 1986.

Dreyfus, H. Why Heideggerian AI Failed and How Fixing it Would Require Making it More Heideggerian. Philosophical Psychology, 20:2, 247-268. 2007.

Dicarlo, Fox, TRENDS in Cognitive Sciences Vol.11 No.8. 2007.

Edelman, Gerald M. The remembered present: A biological theory of consciousness. New York: Basic Books. 1989.

Edelman, Gerald M, Tononi, Giulio. Bilincin Evreni. Küre yayınları. 2019.

Elman. J.L. Finding Structure in Time. Cognitive Science 14, 179-211 1990.

Fodor., Pylyshyn. Cognition, Voluma 28. Issues 1-1, 1988, pp: 3-71.

Fodor. The Language of Thought. Cambridge, MA; Harvard University Press. 1975.

Fodor. Psychosemantics: The problem of Meaning in the Philosophy of Mind. Cambridge MA: MIT Press. 1987.

Fodor. The elm and the Expert: Mentalese and its semantics. Cambridge MA: MIT Press. 1994.

Fukushima, Neocognition: A self organizing Neural Network Model for a Mechanism of Pattern Recognition Unaffected by Shift in Position. Biol. Cybernetics 36, 193-202. 1980.

Friston. K. The History of the Future of the Bayesian Brain. NeuroImage 62, 1230-1233. 2012.

Goldstein, E.B., Sensation and Perception. Cengage Learning, 2010.

Harlow, H. F. (1949). The formation of learning sets. *Psychological Review, 56*(1), 51-65

Hassabis. D., Kumaran. D., Summerfield. C., Botvinick. M. Neuroscience inspired Artificial Intelligence. Neuron 95 Elviser inc. 2017.

Hawkins, Lewis, Klukas, Purdy, Ahmad. A framework for Intelligence and Cortical Functain Based on Grid Cells in the Neocortex. Front. Neural Circuits. Numenta, Inc., Redwood City, CA, United States. 2019.

Hebb, Donald. The Organization of Behavior. New York: Wiley. 1949.

Hinton, G.E., Dayan, P. Frey. B.J, Neal. R.M. The Wake Sleep Algorithm for Unsupervised Neural Networks . Science. Vol 268. 1995.

Hinton G.E., Osindero S., Teh Y. A fast learning algorithma for deep belief nets. Neural Computatiın 18, 1527-1554. 2006.

Hinton, G.E., Sutskever. I. Deep Narrow Sigmoid Belief Networks are Universal Approximators. Neural Computation 20, 2629-2636. 2008.

Hinton, G. Boltzmann Machines. Encyclopedia of Machine Learning anda Data Mining. DOI 10.1007/978-1-14899-7502-7 31-1 Q Springer Science+Business Media New York. 2014.

Hochreiter. S., Schmidbuber. J. Long short-Term Memory. Neural Computation 9(8): 1735-1780, 1997.

Hopfield. J.J. Neural Networks and Physical System with Emergent Collective Computational Abilities. Proceedings of the National Acedemy of Sciences. Vol. 79, pp: 2554-2558. 1982.

Hopfield. J.J., Feinstein. D.I., Palmer. R.G. Unlearning has a stabilizing effect in collective memory. Nature 304, 158-159. 1983.

Hubel, D. H, Wiesel, T. N. Receptive Fields, Binocular Interaction and Functional Architecture in the Cat’s Visual Cortex. J.Physiol. 160, pp. 106-154, 1962.

Jackson, Frank ‘What Mary Didn’t Know’, Journal of Philosophy, 83, pp. 291–5.

Kahneman. D. Thinking Fast and Slow. Allen Lane Pub. 2011.

Kim, J. (1984), SUPERVENIENCE AND SUPERVENIENT CAUSATION. The Southern Journal of Philosophy, 22: 45-56.

Kim, Jaegwon, 1998, *Mind in a Physical World*, Cambridge, MA: MIT Press.

Koch, C. The Quest for Consciousness. Roberts&Company Pub. 2004.

Koch, C. Consciousness, Confessions of a romantic reductionist-The MIT Press. 2012.

Koch, Christof, Tononi, Giulio; Boly, Melanie; Massimini, Marcello. "Integrated information theory: from consciousness to its physical substrate". Nature Reviews Neuroscience. 17 (7): 450–461. 2016.

Kosslyn, S., Ganis, G., Thompson, W. Neural Foundations of İmagery. Nature Reviews, Neuroscience. Vol: 2, pp: 635-642. 2001.

Lake. B., Ullman. T., Tenenbaum, J., &Gershman, S. (2017). Building machines that learn and think like people. Behavioral and brain Sciences, 40, E253.

Lakoff. G. The Neural Theory of Methapor. Report: Jaunary 2009. An earlier version appeared in: R.Gibbs. 2008 The methapor Hnadbook, Cambridge University Press.

Lakoff. G. Explaining Embodied Cognition Results. Topics in Cognitive Science 4 773-785. 2012.

Lakoff. G. Mapping the brain’s metaphor cicuitry: metaphorical thought in everyday reason. Frontiers in Neuroscience. Vol.8, Arcticle 958. 2014.

LeCun, Boser, Denker, Henderson, Howard, Hubbard, Jackel. Backpropagation applied to handwritten zip code recognition. Neural Computation 1-4 541-6551. 1989.

LeCun, Bengio, Hinton. Deep Learning. Nature 521 (7553), 436-444. 2015.

Levine, Joseph. ‘Materialism and Qualia: The Explanatory Gap’, Pacific Philosophical Quarterly, 64, pp. 354–61. 1983.

Levin, Janet, "Functionalism", The Stanford Encyclopedia of Philosophy (Winter 2021 Edition).

Lewis. D., (1972): Psychophysical and theoretical identifications, Australasian Journal of Philosophy, 50:3, 249-258

Lewis. D., **Mad pain** and Martian **pain.** In Ned Block (ed.), Readings in the Philosophy of Psychology. Harvard University Press. pp. 216-222 1980.

Libet. B. Do we have free will?. Journal of Consciousness Studies, 6, No. 8-9, pp. 47-57. 1999.

Luo Liqun. Principles of Neurobiology. Garland Science, Taylor & Francis Group, LLC. 2016.

Lycan, W. (1991). Homuncular Functionalism Meet PDP. In Ramsey, Stich and Rumelhart, eds., Philosophy and Connectionist Theory. L. Erlbaum: 259–86.

McCarthy, J. Programs with common sense. Symposium on The Mechanization of

Thought Processes. 1959.

McClelland, Thomas. Connectionist Models of Cognition. Cambridge handbook on computationanl cognitive modelling. Cambridge University Press. 2008.

Mclever. M., Schmitz. L., Mugan. U., Murphey. T., Mobley. C.D. Massive increase in visual range preceded the origin of terrestial vertebrates. PNAC 114(12), E2375-E2384. 2017.

McCulloch, Warren; Walter Pitts. A Logical Calculus of Ideas Immanent in Nervous Activity. Bulletin of Mathematical Biophysics. 5 (4):115–133. 1943.

Minsky, Marvin. Steps Toward Artificial Intelligence, Proc. IRE, 49, 1, pp. 8-30. 1961.

Minsky, Marvin. Papert, Seymour. Perceptrons. Maple Press Company, M.I.T . 1969.

Minsky, Marvin. A Framework for Representing Knowledge. MIT-AI Laboratory Memo 306, June, 1974.

Mnih, V., Kavukcuoglu, K., Silver, D. et al. Human-level control through deep reinforcement learning. Nature 518, 529–533 (2015).

Mountcastle, V. An Organizing Principle for Cerebral Function: The Unit Model and the Distributed System. The Mindful Brain (Gerald M. Edelman and Vernon B. Mountcastle, eds.) Cambridge, Massachusetts: MIT Press. 1978.

Mountcastle, V. The columnar organization of the neocortex. Review Article. Brain, 120, 701-722. 1997.

Moser, May-Brit, Rowland. David C. Moder Edward. Place cells, Grid Cells, and

Memory Cold Spring Harbor Perspectives in Biology. 7:a021808. 2015.

Nagel, T. “Yarasa Olmak Nasıl Bir Şeydir?”(1974), Nagel. T. Zihin ve Evren. Jaguar kitap. 2015.

Neal. Radford. Connectionist learning of belief networks. Artificial Intelligence 56: 71-113. 1992.

Newell, Alan and Simon, Herbert A. Human problem solving. Engle- wood Cliffs, NJ: Prentice Hall. 1972.

Nilsson, Nils Yapay Zeka. Boğziçi Üniversitesi Yayınevi, 2011.

Penrose, Roger. Kralın Yeni Aklı, Koç Üniversitesi Yayınları, 2015.

PLACE, U. T., Is consciousness a brain process , British Journal of Psychology, 47:1 (1956) p.44

Pinker, Steven. Zihin Nasıl Çalışır. Alfa Basım, 2016.

Pinker, Steven. Dil İçgüdüsü. Bilge Kültür Sanat, 2018.

Alexander Pritzel, Benigno Uria, Sriram Srinivasan, Adrià Puigdomènech Badia, Oriol Vinyals, Demis Hassabis, Daan Wierstra, Charles Blundell . Neural Episodic Control. Proceedings of the 34th International Conference on Machine Learning, PMLR 70:2827-2836, 2017.

Putnam, Hilary. Nature of mental states. Intentionality, Minds and Perception, ed. Hector-Neri Castañeda Detroit, MI: Wayne State University Press, 1967.

Rosch. E. Principles of Categorization. in Rosch, E. & Lloyd, B.B. (eds), Cognition and Categorization, Lawrence Erlbaum Associates, Publishers, (Hillsdale), 1978.

Rosenblatt, F. "The Perceptron: A Probabilistic Model For Information Storage And Organization In The Brain". Psychological Review. 65 (6): 386–408. 19 1958.

Ryle. The Concept of Mind. Londra:Hutchingson. 1949.

Searle, John. R. Minds, Brains, and Programs. Behavioral and Brain Sciences 3(3): 417-457. 1980.

Seth, A. Being You, a New Science of Consciousness. Faber epub. 2021.

Shanahan, Murray, "The Frame Problem", The Stanford Encyclopedia of Philosophy (Spring 2016 Edition), 2009.

Smolensky. On the Proper Treatment of Connectionism. Behavioral and brain sciences (1988). 11, 1-74.

Sutton, Barto. Reinforcement Learning: An Introduction. MIT Press, Cambridge, MA, A Bradford Book, 1998.

Sutton, R. S., & Barto, A. G. (1990). Time-derivative models of Pavlovian reinforcement. In M. Gabriel & J. Moore (Eds.), Learning and computational neuroscience: Foundations of adaptive networks (pp. 497–537). The MIT Press.

Schultz., Hollerman. Dopamine neurons report an error in the temporal prediction of reward during learning. Nat Neurosci 1, 304–309 (1998).

Teyler, Discenna. Behavioural Neuroscience 100(2): 147-54. 1986.

Teyler, T.J., Rudy, J.W. The hippocampal Indexing Theory and Episodic Memory: Updating the İndex. Hippocampus: commentary. 2007.

Tolman Edward. Cognitive Maps in Rats and Men. The Pyschological Review, 55(4), 189-208. 1948.

Tonini, Koch. Cosciousness: here, there and everywhere?. Phil. Trans. R. Soc. B 370: 20140167. 2015.

Tonini, Massimini, Boly, Koch. İntegrated Information Theory: From Consciousness to its physical substrate. Nature Reviews Neuroscience. Vol: 17 (7): 450-461. 2016.

Turing, Alan. On Computable Numbers with Applications to the Entscheidung-Problem. Proceedings of the London Mathematical Society, 42: 230–265. 1936.

Turing, A.M. (1948), 'Intelligent Machinery', National Physical Laboratory Report, in B. Meltzer and D. Michie, eds, Machine Intelligence 5, Edinburgh University Press (1969).

Turing, Alan. Computing Machinery and Intelligence. Mind, New Series, Vol. 59, No. 236 (Oct., 1950), pp. 433-460. 1950.

Yamins. D., DiCarlo J.J. Using goal driven deep learning models to understand sensory cortex. Nature neuroscience vol 19-3. 2016.

Young, R., Lewis, R., The Soar Cognitive Architecture and Human Working Memory. Cambridge University Press. 1999.

Wang, J.X., Kurth-Nelson, Z., Kumaran, D. et al. Prefrontal cortex as a meta-reinforcement learning system. Nat Neurosci 21, 860–868 (2018).

Williams, D. E., Hinton G. E., & Williams R. J.  “Learning Representations by Back-propagating Errors,” Nature, 323.10: 533–536. 1986.



# EK 1. ORİJİNALLİK RAPORU

HACETTEPE ÜNİVERSİTESİ
SOSYAL BİLİMLER ENSTİTÜSÜ
YÜKSEK LİSANS TEZ ÇALIŞMASI ORİJİNALLİK RAPORU

HACETTEPE ÜNİVERSİTESİ
SOSYAL BİLİMLER ENSTİTÜSÜ
FELSEFE ANABİLİM DALI BAŞKANLIĞI'NA

Tarih:11/4/2022

Tez Başlığı : Daniel Dennett'ın Zihin Felsefesinde Güçlü Yapay Zekanın Olasılığı

Yukarıda başlığı gösterilen tez çalışmamın a) Kapak sayfası, b) Giriş, c) Ana bölümler ve d) Sonuç kısımlarından oluşan toplam 154 sayfalık kısmına ilişkin, 30/3/2022 tarihinde şahsım/tez danışmanım tarafından Turnitin adlı intihal tespit programından aşağıda işaretlenmiş filtrelemeler uygulanarak alınmış olan orijinallik raporuna göre, tezimin benzerlik oranı % 4'tür.

Uygulanan filtrelemeler:

1- ☐ Kabul/Onay ve Bildirim sayfaları hariç
2- ☐ Kaynakça hariç
3- ☐ Alıntılar hariç
4- ☒ Alıntılar dâhil
5- ☐ 5 kelimeden daha az örtüşme içeren metin kısımları hariç

Hacettepe Üniversitesi Sosyal Bilimler Enstitüsü Tez Çalışması Orijinallik Raporu Alınması ve Kullanılması Uygulama Esasları'nı inceledim ve bu Uygulama Esasları'nda belirtilen azami benzerlik oranlarına göre tez çalışmamın herhangi bir intihal içermediğini; aksinin tespit edileceği muhtemel durumda doğabilecek her türlü hukuki sorumluluğu kabul ettiğimi ve yukarıda vermiş olduğum bilgilerin doğru olduğunu beyan ederim.

Gereğini saygılarımla arz ederim.

11.4.2022

**Adı Soyadı:** Ayşe Çağıl

**Öğrenci No:** N15222033

**Anabilim Dalı:** Felsefe

**Programı:** Felsefe

**DANIŞMAN ONAYI**

UYGUNDUR.

Prof. Dr. Nazile KALAYCI



**HACETTEPE UNIVERSITY**
**GRADUATE SCHOOL OF SOCIAL SCIENCES**
**MASTER'S THESIS ORIGINALITY REPORT**

**HACETTEPE UNIVERSITY**
**GRADUATE SCHOOL OF SOCIAL SCIENCES**
**PHILOSOPHY DEPARTMENT**

11/4/2022

Thesis Title : Possibility of strong artificial intelligence in Daniel Dennett's philosophy of mind

According to the originality report obtained by myself/my thesis advisor by using the Turnitin plagiarism detection software and by applying the filtering options checked below on 30/3/2022 for the total of 154 pages including the a) Title Page, b) Introduction, c) Main Chapters, and d) Conclusion sections of my thesis entitled as above, the similarity index of my thesis is 4 %.

Filtering options applied:

1. ☐ Approval and Decleration sections excluded
2. ☐ Bibliography/Works Cited excluded
3. ☐ Quotes excluded
4. ☒ Quotes included
5. ☐ Match size up to 5 words excluded

I declare that I have carefully read Hacettepe University Graduate School of Social Sciences Guidelines for Obtaining and Using Thesis Originality Reports; that according to the maximum similarity index values specified in the Guidelines, my thesis does not include any form of plagiarism; that in any future detection of possible infringement of the regulations I accept all legal responsibility; and that all the information I have provided is correct to the best of my knowledge.

I respectfully submit this for approval.

11/4/2022

**Name Surname:** Ayşe Çağıl

**Student No:** N15222033

**Department:** Philosophy

**Program:** Philosophy

**ADVISOR APPROVAL**

APPROVED.

Prof. Dr. Nazile KALAYCI

# EK 2. ETİK KURUL / KOMİSYON İZNİ YA DA MUAFİYET FORMU

HACETTEPE ÜNİVERSİTESİ
SOSYAL BİLİMLER ENSTİTÜSÜ
TEZ ÇALIŞMASI ETİK KOMİSYON MUAFİYETİ FORMU

HACETTEPE ÜNİVERSİTESİ
SOSYAL BİLİMLER ENSTİTÜSÜ
FELSEFE ANABİLİM DALI BAŞKANLIĞI'NA

Tarih: 11/4/2022

Tez Başlığı: Daniel Dennett'ın zihin felsefesinde güçlü yapay zekanın olasılığı

Yukarıda başlığı gösterilen tez çalışmam:

1. İnsan ve hayvan üzerinde deney niteliği taşımamaktadır,
2. Biyolojik materyal (kan, idrar vb. biyolojik sıvılar ve numuneler) kullanılmasını gerektirmemektedir.
3. Beden bütünlüğüne müdahale içermemektedir.
4. Gözlemsel ve betimsel araştırma (anket, mülakat, ölçek/skala çalışmaları, dosya taramaları, veri kaynakları taraması, sistem-model geliştirme çalışmaları) niteliğinde değildir.

Hacettepe Üniversitesi Etik Kurullar ve Komisyonlarının Yönergelerini inceledim ve bunlara göre tez çalışmamın yürütülebilmesi için herhangi bir Etik Kurul/Komisyon'dan izin alınmasına gerek olmadığını; aksi durumda doğabilecek her türlü hukuki sorumluluğu kabul ettiğimi ve yukarıda vermiş olduğum bilgilerin doğru olduğunu beyan ederim.

Gereğini saygılarımla arz ederim.

11/4/2022

**Adı Soyadı:** Ayşe Çağıl

**Öğrenci No:** N15222033

**Anabilim Dalı:** Felsefe

**Programı:** Felsefe

**Statüsü:** ☐ Yüksek Lisans ☐ Doktora ☐ Bütünleşik Doktora

**DANIŞMAN GÖRÜŞÜ VE ONAYI**

Prof. Dr. Nazile Kalaycı

HACETTEPE UNIVERSITY
GRADUATE SCHOOL OF SOCIAL SCIENCES

## ETHICS COMMISSION FORM FOR THESIS

**HACETTEPE UNIVERSITY**
**GRADUATE SCHOOL OF SOCIAL SCIENCES**
**PHILOSOHY DEPARTMENT**

11/4/2022

Thesis Title: Possibility of strong artificial intelligence in Daniel Dennett's philosophy of mind

My thesis work related to the title above:

1. Does not perform experimentation on animals or people.
2. Does not necessitate the use of biological material (blood, urine, biological fluids and samples, etc.).
3. Does not involve any interference of the body's integrity.
4. Is not based on observational and descriptive research (survey, interview, measures/scales, data scanning, system-model development).

I declare, I have carefully read Hacettepe University's Ethics Regulations and the Commission's Guidelines, and in order to proceed with my thesis according to these regulations I do not have to get permission from the Ethics Board/Commission for anything; in any infringement of the regulations I accept all legal responsibility and I declare that all the information I have provided is true.

I respectfully submit this for approval.

11/4/2022

**Name Surname:** Ayşe Çağıl

**Student No:** N15222033

**Department:** Philosophy

**Program:** Philosophy

**Status:** ☐ MA ☐ Ph.D. ☐ Combined MA/ Ph.D.

**ADVISER COMMENTS AND APPROVAL**

Prof. Dr. Nazile Kalaycı